

GRADIENT BUNDLE ANALYSIS: A FULL  
TOPOLOGICAL APPROACH TO  
CHEMICAL BONDING

by

Amanda Morgenstern

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Applied Chemistry).

Golden, Colorado

Date \_\_\_\_\_

Signed: \_\_\_\_\_

Amanda Morgenstern

Signed: \_\_\_\_\_

Dr. Mark E. Eberhart  
Thesis Advisor

Golden, Colorado

Date \_\_\_\_\_

Signed: \_\_\_\_\_

Dr. David Wu  
Professor and Head  
Department of Chemistry and Geochemistry

## ABSTRACT

The “chemical bond” is a central concept in molecular sciences, but there is no consensus as to what a bond actually is. Therefore, a variety of bonding models have been developed, each defining the structure of molecules in a different manner with the goal of explaining and predicting chemical properties. While many twentieth century bonding models provide useful information for a variety of chemical systems, these models are sometimes less insightful for more lofty goals such as designing metalloenzymes. The design process of novel catalysts could be improved if more predictive and accurate models of chemical bonding are created. One recently developed bonding model based on the topology of the electron charge density is the quantum theory of atoms in molecules (QTAIM). QTAIM defines bonding interactions as one-dimensional ridges of electron density,  $\rho(\mathbf{r})$ , which are known as bond paths. As with any bonding model, there are instances where bond paths do not adequately describe properties of interest, such as in the analysis of histone deacetylase.

This thesis describes the initial development of gradient bundle analysis (GBA), a chemical bonding model that creates a higher resolution picture of chemical interactions within the charge density framework. GBA is based on concepts from QTAIM, but uses a more complete picture of the topology and geometry of  $\rho(\mathbf{r})$  to understand and predict bonding interactions. Gradient bundles are defined as volumes bounded by zero-flux surfaces (ZFSs) in the gradient of the charge density with well-defined energies. The structure of gradient bundles provides an avenue for detecting the locations of valence electrons, which correspond to reactive regions in a molecule. The number of electrons in bonding regions, which are defined by the lowering of kinetic energy in gradient bundles, is found to correlate to bond dissociation energy in diatomic molecules. Furthermore, site reactivity can be understood and predicted by observing the motion of ZFSs bounding gradient bundles and calculating condensed gradient bundle Fukui functions. Using only the ground state charge density, I

present preliminary results for a method that predicts which regions in a molecule are most likely to undergo nucleophilic and electrophilic attack, effectively locating the HOMO and LUMO.



## TABLE OF CONTENTS

ABSTRACT .....	iii
LIST OF FIGURES .....	viii
LIST OF TABLES .....	xiii
LIST OF SYMBOLS .....	xiv
LIST OF ABBREVIATIONS.....	xvii
CHAPTER 1 INTRODUCTION—TOPOLOGICAL CHEMICAL BONDING MODELS .....	1
1.1 Quantum Theory of Atoms in Molecules.....	3
1.1.1 Critical Points .....	4
1.1.2 Bond Paths .....	7
1.1.3 Atomic Basins and The Zero-Flux Surface Condition .....	9
1.1.4 Delocalization Index.....	11
1.2 Topological Models of Kinetic Energy Density .....	12
1.3 Bond Bundles .....	16
CHAPTER 2 IN SEARCH OF AN INTRINSIC CHEMICAL BOND .....	20
2.1 Abstract .....	20
2.2 Introduction .....	21
2.3 Gradient Bundles .....	23
2.4 Methods .....	25
2.5 Results and Discussion .....	28

2.6	Conclusions.....	33
CHAPTER 3 BOND DISSOCIATION ENERGIES FROM GRADIENT BUNDLE ANALYSIS .....		
		35
3.1	Abstract .....	35
3.2	Introduction .....	35
3.3	Methods .....	38
3.4	Results and Discussion .....	38
3.5	Conclusions.....	41
CHAPTER 4 THE INFLUENCE OF ZERO-FLUX SURFACE MOTION ON CHEMICAL REACTIVITY .....		
		44
4.1	Abstract .....	44
4.2	Introduction .....	45
4.3	Background .....	46
4.4	Derivation of ZFS Influence on Chemical Reactivity .....	50
4.5	Methods .....	54
4.6	Results and Discussion .....	55
4.7	Conclusions.....	60
CHAPTER 5 APPLICATIONS TO ENZYME DESIGN .....		
		62
5.1	QM/DMD Methods .....	63
5.2	Predictive Methods for Computational Metalloenzyme Redesign - A Test Case with Carboxypeptidase A .....	64
5.2.1	Introduction.....	64
5.2.2	Methods .....	66
5.2.3	Results and Discussion .....	67

5.3	Metal Specificity of Histone Deacetylase .....	71
5.3.1	Introduction.....	71
5.3.2	Methods .....	74
5.3.3	Results and Discussion .....	78
5.4	Conclusions.....	83
CHAPTER 6 FUTURE WORK AND CONCLUSIONS .....		84
6.1	Gradient Bundle Condensed Fukui Functions .....	85
6.2	Predicting Site Reactivity from the Ground State Charge Density .....	91
6.3	Gradient Bundle Analysis Method Improvements.....	96
6.4	Applications to Aromaticity .....	97
6.5	Gradient Bundle Analysis in Closed Systems.....	99
6.6	Electrostatic Preorganization and Active Site Model Size .....	100
6.7	Conclusions.....	102
REFERENCES CITED .....		104
APPENDIX A FUNCTIONAL DEPENDENCE .....		122
APPENDIX B HDAC8 METAL SWAPPING DATA .....		123
APPENDIX C PERMISSIONS .....		125
C.1	Journal Permissions .....	125
C.2	Author Permissions .....	125

## LIST OF FIGURES

Figure 1.1	The CPs and bond paths in a cubane molecule. The image is colored as follows: C-grey, H-white, bond CPs-red, ring CPs-green, cage CPs-blue. Critical points and bond paths were calculated using the Amsterdam Density Functional package . . . . .	5
Figure 1.2	Bond paths in cyclopropane. The outward curved bond paths between the carbon atoms (colored lines) indicate ring strain in the molecule. The internuclear axes (grey lines) are shown for comparison. . . . .	8
Figure 1.3	A contour plot of ELF for a Ge <sub>2</sub> molecule showing topological separation of core, valence, and bonding regions into separate regions bounded by ZFSs. . . . .	13
Figure 1.4	2D cut plane of the charge density of ethene in the plane of the molecule. Thick colored lines show a sampling of ZFSs bounding the C-C bond CP. . . . .	17
Figure 2.1	A C-C bond bundle in ethene built up from its irreducible bundles. The black lines shown in the top left are the 1-ridges that form the edges of one IB that is a part of the C-C bond bundle in ethene. The 3D IB bounded by ZFSs is shown in the bottom left, and the full bond bundle composed of the eight symmetry equivalent IBs is on the right. . . . .	23
Figure 2.2	A gradient bundle created from the triangulation of a sphere (shown in yellow) around a carbon nucleus in benzene. Carbon nuclei are shown in black, hydrogen nuclei in white, bond critical points in cyan, and bond paths are black. . . . .	24
Figure 2.3	Gradient paths seeded every 10° in the charge density for N <sub>2</sub> and Ne <sub>2</sub> to show the general gradient path behavior. Dashed lines represent contour lines of $\rho(\mathbf{r})$ , solid lines are gradient paths, blue circles are nuclear CPs, cyan circles are bond CPs, and red lines are bond paths. . . . .	25
Figure 2.4	In all plots the gradient bundles are numbered in a counterclockwise manner with the first GB containing the bond path and the 72 <sup>nd</sup> GB containing the ridge from the nuclear CP in the direction opposite the bond path. Shown for clarity are gradient paths used to create ZFSs every 20°. . . . .	27

Figure 2.5	A representation of a gradient bundle in $\text{Ne}_2$ bounded by ZFSs at $\theta = 40^\circ$ and $\theta = 50^\circ$ that has been rotated around the internuclear axis by $280^\circ$ . The image has been tilted to emphasize the structure of a gradient bundle which takes the shape of an umbrella. Note that the gradient bundles in this system will actually extend out to infinity; the GB shown here has been truncated in this image for clarity. ....	28
Figure 2.6	Electron and kinetic energy integrations over gradient bundles for a set of homonuclear diatomic molecules. Left: The electron count in each gradient bundle. Right: Kinetic energy density in each gradient bundle in electron volts (eV). Note that the scales on the plots for hydrogen vary from the rest of the dimers. ....	30
Figure 2.7	Average kinetic energy per electron in each gradient bundle. Note that the scales on the plot for hydrogen vary from the rest of the dimers. ....	31
Figure 3.1	A plot of the relationship between total energy, kinetic energy, and potential energy in a homonuclear diatomic molecule using the definitions from eqn 3.1. The total energy was calculated using a pair potential of the form $-a/r^5 + b/r^9$ . ....	37
Figure 3.2	Integration of the average KE/e in each $2.5^\circ$ gradient bundle in CO. The plots are oriented so that they line up geometrically with Figure 3.3: the bonding region is in the center of the image ( $0^\circ$ ) and the lone pair regions are on the outside edges of the plots ( $180^\circ$ ). ....	39
Figure 3.3	A cut plane of the kinetic energy bonding region in carbon monoxide. Contour lines of $\rho(\mathbf{r})$ are drawn on a logarithmic scale from $10^{-3} - 10$ e/bohr <sup>3</sup> . Thick black lines show bond paths and interatomic surfaces. Thin black lines show gradient paths seeded every $10^\circ$ around each nuclei. The red shaded region is a 2D cut plane of half of the KE BR. ....	40
Figure 3.4	Relationship between the normalized electron count in KE BRs and BDE. Experimental BDEs were obtained from .....	42
Figure 4.1	A rotational gradient bundle contained in the iodine atomic basin of an ICCF molecule. Contours on a cut plane of $\rho(\mathbf{r})$ through the molecule are drawn on a logarithmic scale from $10^{-4} - 10$ e/bohr <sup>3</sup> . Spheres are colored as: I-purple, C-black, F-green, bond CP-cyan; bond paths are grey, and a sampling of gradient paths on the ZFSs are shown in black. This coloring is used in all figures in this chapter. ....	47

Figure 4.2	Atomic surface motion for ICCF $\pm 1$ electron. Left: ICCF anion interatomic surface motion. Right: ICCF cation interatomic surface motion. Red (blue) regions indicate the motion of atomic boundaries for the $N+1$ ( $N-1$ ) molecule, black lines show the neutral molecule interatomic surfaces. Contours of $\rho(\mathbf{r})$ on the cut plane are drawn for the neutral molecule. ....	56
Figure 4.3	Overlays of a cut plane of ZFSs for the top half of a neutral ICCF and [ICCF] <sup>+</sup> (top) and [ICCF] <sup>-</sup> (bottom) molecules. Blue (red) lines represent ZFSs for the $N - 1$ ( $N + 1$ ) molecule and black ZFSs are for the neutral molecule. Contour lines at $10^{-3}$ e/bohr <sup>3</sup> are shown in solid black for the neutral molecule and dashed blue (red) for the cation (anion). Contour flooding of $\rho(\mathbf{r})$ on the cut plane is for the neutral molecule. ....	59
Figure 4.4	The HOMO (left) and LUMO (right) for a neutral ICCF molecule. ....	59
Figure 5.1	Active site of CPA showing partitioning for QM/DMD calculations: grey-QM/DMD region, blue-QM region. Single point DFT calculations were performed on the entire active site pictured here. ....	65
Figure 5.2	The proposed water-promoted mechanism of peptide hydrolysis in CPA (the peptide substrate, shown in blue, is truncated for clarity). The water molecule is shown in red. ....	66
Figure 5.3	Bond paths of interest in the native CPA (A) and V243R mutant (B) ES complexes. Contours in $\rho(\mathbf{r})$ are drawn on a cut plane on a logarithmic scale from $10^{-3} - 1$ e/bohr <sup>3</sup> . Red lines indicate bond paths. The pictured portion of the Zn-O1 bond bundles are shaded green with black lines showing approximate edges. The following coloring scheme is used: Zn-purple, O-red, C-black, H-white, N-blue, bond CP-cyan, ring CP-green. ....	68
Figure 5.4	Bond paths and critical points for the both the native CPA (A) and V243R mutant (B) TS. The ring CP (green) has moved towards the Zn-O <sub>wat</sub> bond CP in both enzymes. After the transition state, the ring CP converges with the Zn-O <sub>wat</sub> bond CP, causing the two critical points to be topologically annihilated and the Zn-O <sub>wat</sub> bonding interaction to break. ....	70

Figure 5.5	The active site of HDAC8 extracted from the crystal structure, highlighting the most critical residues relevant to catalysis within the active site. In this work, F306 was mutated to tyrosine to facilitate catalysis. This full structure was used as the QM region in QM/DMD calculations and for the QM mechanistic study and charge density calculations.....	72
Figure 5.6	Proposed proton shuttle mechanism for HDAC8. ....	73
Figure 5.7	Calculated reaction pathway for HDAC8 with a variety of divalent metal ions. ....	75
Figure 5.8	Schematic of all considered thermodynamic cycles exploited for relative metal binding affinities in HDAC8. ....	76
Figure 5.9	Critical points and bond paths of interest in the active site of HDAC8 with Zn (left) and Mn (right). When Zn is present in the active site a bond path forms between the water oxygen and the carbon atom from the substrate carbonyl. This topologically necessitates a ring critical point to exist in the active site as well. When Mn is present, no ring CP is found as there is not a bond path between the water and substrate carbonyl. Ni, Fe, and Co give the same topology as Zn, while Mg has the same topology as Mn. Sphere coloring is as follows: C-black, N-blue, O-red, Metal-grey, bond CP-cyan, ring CP-green. ....	79
Figure 5.10	Relationship between $\rho(\mathbf{r})$ at $H_{H180}$ -metal bond CPs and $\Delta\Delta G$ of metal swapping. ....	81
Figure 6.1	Plots of $f_{GB}^-$ (left) and $f_{GB,\%}^-$ (right) around one fluorine atom in $F_2$ . Gradient bundle angles along the x-axes are measured with respect to the bond path, such that the GB at $0^\circ$ contains the bond path and is bounded by the interatomic surface. ....	87
Figure 6.2	Plots of $f_{GB}^+$ (left) and $f_{GB,\%}^+$ (right) for one fluorine atom in $F_2$ . ....	87
Figure 6.3	The HOMO (left) and LUMO (right) for a neutral fluorine molecule. Black dashed boxes indicate the regions of the orbital plots that geometrically match the Fukui function plots in Figure 6.1 and Figure 6.2. ....	87
Figure 6.4	Plots of $f_{GB,\%}^-$ (left) and $f_{GB,\%}^+$ (right) around one nitrogen atom in $N_2$ . ..	89
Figure 6.5	The HOMO (left) and LUMO (right) for a neutral nitrogen molecule. Black dashed boxes indicate the regions of the orbital plots that geometrically match the Fukui function plots in Figure 6.4. ....	89

Figure 6.6	Left: A cutplane of rotational gradient bundles in $N_2$ . Black circles indicate points on the bounding zero-flux surfaces at an isovalue of $\rho(\mathbf{r}) = 0.002$ . Right: The distance between GB points on the left, which we refer to as the width of the gradient bundles. The x-axis displays the gradient bundle angle, where $0^\circ$ is the gradient bundle bounded by the interatomic surface and containing the bond path.....	93
Figure 6.7	Left: A cutplane of rotational gradient bundles in $F_2$ . Black circles indicate points on the bounding zero-flux surfaces at an isovalue of $\rho(\mathbf{r}) = 0.002$ . Right: The width of the gradient bundles measured at the points marked in the left plot. ....	94
Figure 6.8	Top: A cut plane of ZFSs bounding rotational gradient bundles in ICCF. The isosurface of $\rho(\mathbf{r}) = 0.002$ is traced with a black line to show where the width of gradient bundles is measured. Middle: The width of each GB at the van der Waals radius. Bottom: $f_{GB,\%}^-$ for each atom in ICCF.....	95
Figure A.1	Average KE/e in gradient bundles seeded every $2.5^\circ$ around a nucleus in $N_2$ using LDA, BLYP, B3LYP, and M06L functionals within the ADF software package. Gradient bundle angle is measured with respect to the bond path. Single point calculations were run using a TZP basis set using the geometry obtained from initial PBE calculations. ....	122



## LIST OF TABLES

Table 2.1	Bond energies and distribution statistics for gradient bundles in homonuclear diatomic molecules. ....	29
Table 3.1	Total number of electrons, number of valence electrons, and normalized electron counts in kinetic energy bonding regions in a set of diatomic molecules. ....	41
Table 4.1	Integration of electrons for the integrals in eqn (4.21) for $[I-C_1-C_2-F] \pm 1$ electron. ....	55
Table 5.1	Bader charges of atoms participating in the tetrahedral intermediate in the CPA hydrolysis mechanism. $O_{\text{wat}}$ is the water oxygen and C1 and O1 are the carbon and oxygen atoms on the substrate carbonyl. ....	69
Table 5.2	Charge density at bond and ring critical points pictured in Figure 5.3 for the rate-determining step of peptide hydrolysis in native CPA and the V243R mutant. ....	69
Table 5.3	$\Delta\Delta G$ of binding between metal ions and HDAC8, relative to $Co^{2+}$ . Energies are in <i>kcal/mol</i> . ....	77
Table 5.4	Predicted and experimental total catalytic activities for different metal ions in HDAC8. Binding affinity energies were taken relative to $Co^{2+}$ . Normalized values are relative to the predicted most active metal ion-protein complex. ....	82
Table B.1	List of literature binding energies with regards to a series of chelating agents for all relevant metal ions studied in HDAC8. ....	123
Table B.2	Calculated $\Delta G$ values for DTPA metal swapped with GEDTA ( <i>kcal/mol</i> )..	124
Table B.3	Experimental $\Delta G$ values for DTPA metal swapped with GEDTA ( <i>kcal/mol</i> ). ....	124
Table B.4	Calculated $\Delta G$ – experimental $\Delta G$ values for DTPA metal swapped with GEDTA ( <i>kcal/mol</i> ). ....	124

## LIST OF SYMBOLS

Activation energy .....	$\Delta G^\ddagger$
Atom condensed Fukui function .....	$f_i$
Atomic Basin .....	$\Omega$
Average Directionality .....	$\langle \tan(\theta) \rangle$
Basis functions .....	$\varphi$
Boundary of mononuclear region .....	$\partial\Omega$
Charge density at a bond critical point .....	$\rho_b$
Charge in ELI-D cell .....	$Q_i$
Chemical hardness .....	$\eta$
Chemical potential .....	$\mu$
Correlation kinetic energy .....	$T_c$
Curvature of charge density in principle direction $i$ .....	$\lambda_i$
Delocalization index .....	$\delta$
Density matrix .....	$P$
Diagonalized Hessian .....	$\Lambda$
Eigenvectors of the Hessian .....	$e_i$
Electron charge density .....	$\rho(\mathbf{r})$
Electronegativity .....	$\chi$
Electrons count .....	$N$
Ellipticity .....	$\varepsilon$

Equilibrium distance .....	$r_{eq}$
Fermi correlation .....	F
Fukui eigenvalues .....	$\mathbf{F}_{ii}$
Fukui function .....	$f(\mathbf{r})$
Fukui matrix .....	$\mathbf{f}$
Fukui orbitals .....	$\phi_i$
Gibbs free energy .....	$G$
Gradient bundle .....	$\omega$
Gradient bundle condensed Fukui function .....	$f_{GB}$
Gradient kinetic energy .....	G
Laplacian .....	$\nabla^2 \rho(\mathbf{r})$
Localized orbital locator .....	$\nu(\mathbf{r})$
Non-interacting kinetic energy .....	$T_s$
Normal vector .....	$\mathbf{n}(\mathbf{r})$
Nuclear coordinates .....	$X_A$
Number of electrons .....	$n_e$
One-electron operator .....	$\hat{O}$
Overlap integral .....	$S$
Percent contribution of ZFS motion to electron response .....	$\% \Delta N_{ZFS}$
Positive kinetic energy .....	$\tau$
Schrödinger kinetic energy .....	K
Total energy .....	$E$
Total kinetic energy .....	$T$

Total potential energy .....	$V$
Uniform electron gas kinetic energy .....	$D_0$
Valence electron density .....	$\rho_V$
Wave function .....	$\psi(\mathbf{r})$
Work of separation .....	$W_\infty$

## LIST OF ABBREVIATIONS

Amsterdam density functional package .....	ADF
Atomic orbital .....	AO
Becke-3-Lee-Yang-Parr .....	B3LYP
Becke-Lee-Yang-Parr .....	BLYP
Bond bundle .....	BB
Bond dissociation energy.....	BDE
Bond order .....	BO
Born-Oppenheimer molecular dynamics .....	BOMD
Carboxypeptidase A.....	CPA
Conceptual density functional theory .....	CDFT
Conductor-like screening model .....	COSMO
Critical point .....	CP
Delocalization Index.....	DI
Density functional theory .....	DFT
Discrete molecular dynamics .....	DMD
Double zeta polarized .....	DZP
Electron .....	e
Electron localizability indicator .....	ELI
Electron localization function.....	ELF
Electron localization index D .....	ELI-D

Electron preceding picture .....	EPP
Enzyme-intermediate .....	EI
Enzyme-substrate .....	ES
Experimental bond energy .....	EBE
Face-centered cubic .....	FCC
Fragments of molecular response .....	FMR
Generalized gradient approximation .....	GGA
Gradient bundle .....	GB
Gradient bundle analysis.....	GBA
Harmonic oscillator model of aromaticity.....	HOMA
Highest occupied molecular orbital.....	HOMO
Hippuryl-L-phenylalanine .....	HPA
Histone deacetylase .....	HDAC
Interatomic surface .....	IAS
Irreducible bundle .....	IB
Kinetic energy .....	KE
Kinetic energy bonding region.....	KE BR
Local density approximation.....	LDA
Localized orbital locator .....	LOL
Lowest unoccupied molecular orbital .....	LUMO
Minnesota 06 local .....	M06L
Molecular dynamics .....	MD
Nucleus-independent chemical shift .....	NICS

Partial electron localization index D .....	pELI-D
Perdew-Burke-Ernzerhof .....	PBE
Quadruple zeta 4 polarized .....	QZ4P
Quantum mechanics .....	QM
Quantum theory of atoms in molecules .....	QTAIM
Reactant state destabilization .....	RSD
Redox induced electron transfer .....	RIET
Response of molecular fragments .....	RMF
Root mean square deviation .....	RMSD
Tao Perdew Staroverov Scuseria .....	TPSS
Theoretical bond energy .....	TBE
Transition state .....	TS
Transition state stabilization .....	TSS
Triple zeta polarized .....	TZP
Valence shell electron pair repulsion .....	VSEPR
Zero-flux surface .....	ZFS
Zeroth order regular approximation .....	ZORA

# CHAPTER 1

## INTRODUCTION—TOPOLOGICAL CHEMICAL BONDING MODELS

The evolution of chemical bonding models and creation of entirely new models is imperative for the progression of the field of chemistry, cultivating a better understanding and control of interactions between atoms. Most models of the chemical bond used today are rooted in century old ideas developed by Lewis, London, and Heitler. Few foundational changes have been made to the way we think about chemical bonds since the work of Pauling, Mulliken, and Slater in the 1930s. While some unique models have been proposed, such as the quantum theory of atoms in molecules (QTAIM) [1] and conceptual density functional theory (CDFT) reactivity indicators [2], these viewpoints are not widely utilized and have yet to find their way into most introductory chemistry texts. By radically changing our view of atomic interactions there is potential to revolutionize how we design new molecules and materials.

One area where an innovative view of chemical bonding could be especially useful is in the design of novel proteins. The design of new enzymes to catalyze non-native reactions can lead to more environmentally friendly syntheses and increased production of medicines and materials. Enzyme catalysis is not a fully understood process, and using computational modeling to simulate the affect of mutations to enzymes is a difficult task [3–6]. Even newer chemical bonding models such as QTAIM can not always properly explain the way chemical bonds will transform in designed enzymes, as is shown in chapter 5.

This thesis presents the foundational work in the development of gradient bundle analysis (GBA). GBA is a potentially transformative bonding model that is based on quantum mechanics, recovers observed properties of chemicals, and is predictive in nature. Gradient bundles are defined using the gradient field of the electron charge density,  $\rho(\mathbf{r})$ , and provide a higher resolution picture of bonding interactions than was previously possible with standard



QTAIM methods. The overarching goal of this work is to provide a useful bonding model that is accessible and applicable to chemists specializing in any field. Two areas where we are currently working on applying GBA are in enzyme design and to complex metallurgical problems such as understanding brittle failure in iridium (see section 6.5).

The creation of any new model necessarily begins by demonstrating its ability to recover known properties in simple molecules. To this end, the first test of GBA is presented in chapter 2, where the locations of valence electrons are recovered using values calculated directly from the ground state charge density. Valence shell electron pair repulsion (VSEPR) diagrams for a set of homonuclear diatomic molecules are reproduced by partitioning the charge density and kinetic energy into gradient bundles, which are volumes bounded by zero-flux surfaces (ZFSs) in the gradient of the charge density,  $\nabla\rho(\mathbf{r})$ . The usefulness of this structure is elucidated in Chapter 3, with the discovery of a quantitative relationship between experimentally determined bond dissociation energies of diatomic molecules and the lowering of kinetic energy in bonding regions defined using gradient bundles.

Once the framework of GBA is in place, we can begin the search for answers to more complex chemical problems. Chapter 4 presents a method for visualizing chemical reactivity based on the motion of ZFSs in  $\nabla\rho(\mathbf{r})$ . This visualization method provides an alternative to drawing electron pushing arrows or using electron transfer between frontier orbitals to picture chemical reactions occurring. The results from this study provide motivation for a method of answering the open question originally posed by Slater, “where are the HOMO and LUMO in the ground state charge density?”. The first theorem of density functional theory (DFT) states that all properties are determined by the charge density of a system. Therefore, it should be possible to predict where electrophilic or nucleophilic attack will occur in a molecule based purely on the ground state charge density, which should also generally match where the HOMO and LUMO are located in a molecule, respectively. We have begun to answer this question using the size and shape of gradient bundles, and the preliminary results of this study are presented in chapter 6. To put the usefulness and complexity of

GBA into context, the remainder of this chapter briefly reviews topological bonding models, with a focus on identifying specific properties that can be recovered using each definition of chemical structure.

## 1.1 Quantum Theory of Atoms in Molecules

The quantum theory of atoms in molecules uses the topology of the  $\rho(\mathbf{r})$  to define and analyze atomic properties and bonding interactions. The main applications of QTAIM to chemical bonding have been on characterizing types of bonding as well as better understanding the strength of interactions and therefore the stability of molecules. This theory has been applied with much success to a wide variety of chemical systems including small molecules [7, 8], biological systems [9–11], and solids [12–14]. QTAIM was originally developed by Professor Richard Bader in the 1970s. The motivation for this theory is given by the following two questions posed by Bader [1, 15]

1. Does an atom in a molecule exist, and if so, how do you recognize one?
2. Are bonds physical observables and how do you define them?

To discuss the physical existence of something you must have some way of observing the item in question. While wave functions,  $\psi(\mathbf{r})$ , give great insight into chemical interactions, they are not quantum-mechanical observables. They are mathematical constructs that approximate reality and are often complex valued functions. Charge density, on the other hand, is a measurable feature and always a real valued function. With the advancement of x-ray diffraction techniques, chemists are now able to image the electron density. It has been argued that an accurate bonding model must be rooted in something real [16, 17]. This is why the electron density was an obvious starting point for QTAIM. The original theory of QTAIM gives insight into bonding interactions in terms of two topological features: bond critical points in  $\rho(\mathbf{r})$  (and their corresponding bond paths) as well as the electron exchange between atomic basins.

### 1.1.1 Critical Points

Chemical bonding information can often be recovered by examining critical points (CPs) in the charge density. There are four types of CPs in a 3D scalar field such as  $\rho(\mathbf{r})$ : local minima, local maxima, and two types of saddle points. A CP in 3D space is defined as

$$\nabla\rho(\mathbf{r}) = \mathbf{i}\frac{d\rho}{dx} + \mathbf{j}\frac{d\rho}{dy} + \mathbf{k}\frac{d\rho}{dz} \rightarrow \begin{cases} = \mathbf{0} & \text{at critical points and } \infty \\ \neq \mathbf{0} & \text{at all other points.} \end{cases} \quad (1.1)$$

CPs are denoted by the number of dimensions of the space and the net number of positive curvatures. For example, at a minimum, the curvatures in all principal directions are positive; therefore, this is a (3, +3) CP.

The ground state charge density at an atomic nucleus is always a maximum, a (3, −3) CP, hence it is called a nuclear CP (within the coulomb approximation the density is actually cusp at a nuclear CP since the finite size of the nuclei are neglected in quantum mechanical calculations, but the point still acts as a maximum, topologically [18, 19]). The simplest topological connection results from a shared (3, −1) saddle point between two nuclear CPs and is indicative of a charge density ridge originating at the (3, −1) CP and terminating at the nuclear CPs (see Figure 1.1). This ridge possesses the topological properties imagined for a chemical bond and is referred to as a bond path. The (3, −1) CP is thus called a bond CP. Other CPs provide additional information about chemical structure. A (3, +1) CP is required at the center of ring structures (rings of bond paths), earning its designation as a ring CP. Cage structures must enclose a single (3, +3) CP so these points are called cage CPs.

The amount of charge density at a bond CP has been used to determine bond strength [1]. Oftentimes, bond order (BO) can be determined by

$$\text{BO} = e^{A(\rho_b - B)} \quad (1.2)$$

where A and B are constants that are dependent on the bond in question and  $\rho_b$  is the value of the charge density at the bond CP. The correlation between bond strength and electron density at bond CPs has been especially useful in characterizing hydrogen bonds [21].

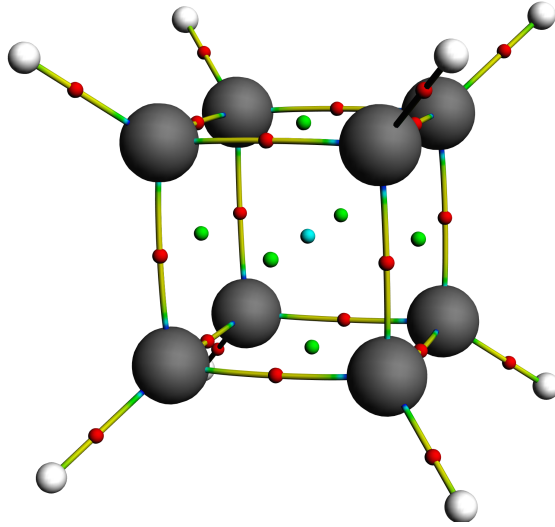


Figure 1.1: The CPs and bond paths in a cubane molecule. The image is colored as follows: C-grey, H-white, bond CPs-red, ring CPs-green, cage CPs-blue. Critical points and bond paths were calculated using the Amsterdam Density Functional package [20].

The Hessian of  $\rho(\mathbf{r})$  determines the curvature at critical points. Its diagonalized form is

$$\Lambda = \begin{bmatrix} \frac{\partial^2 \rho}{\partial x^2} & 0 & 0 \\ 0 & \frac{\partial^2 \rho}{\partial y^2} & 0 \\ 0 & 0 & \frac{\partial^2 \rho}{\partial z^2} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (1.3)$$

where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are the curvatures of the density in 3 principle directions. Its trace, the Laplacian,  $\nabla^2 \rho(\mathbf{r})$ , is the sum of the three diagonal elements of the Hessian and has been used to characterize bonding interactions. At a bond CP, one of the eigenvectors of the Hessian is parallel to the bond path and the other two are perpendicular. Since  $\rho(\mathbf{r})$  is a maximum perpendicular to the bond path at a bond CP, these two eigenvalues ( $\lambda_1$  and  $\lambda_2$ ) are negative while the component parallel to the bond path ( $\lambda_3$ ) is always positive.

Bader argued that a positive value of the Laplacian indicates closed-shell bonding, such as ionic or van der Waals interactions [1]. A positive value of  $\nabla^2 \rho_b$  occurs when the curvature along the bond path is greater than the sum of the curvatures perpendicular, indicating a depletion of charge density along the bond path. On the other hand, a negative value of  $\nabla^2 \rho_b$  is associated with open shell (covalent) bonding since there is an accumulation of charge

density at the bond CP [22].

Grabowski calculated the amount of charge density and the value of the Laplacian at bond critical points in systems with a large variety of hydrogen bonding interactions ranging from extremely strong  $(\text{F}-\text{H}-\text{F})^-$  to extremely weak  $\text{HCCH}-\pi(\text{HCCH})$  interactions in different chemical environments [23]. He found that parameters purely from the proton-donating part of the H-bond interaction (such as  $\rho_b$  and  $\nabla\rho_b$ ) correlate well to hydrogen bond strength, so there is no need to analyze the proton-accepting part of the H-bond or the environment around the bond.

Ford recently performed a study analyzing values at bond critical points for lithium-bonded complexes such as  $\text{LiCl}\cdot\text{NH}_3$  and  $\text{LiF}\cdot\text{H}_2\text{S}$  [24]. He was able to predict which structures formed rings, for example,  $\text{LiF}\cdot\text{H}_2\text{O}$ , from structures that were lacking ring CPs, such as  $\text{LiF}\cdot\text{H}_2\text{S}$ . Correlation was also found between the amount of charge density and the values of the Laplacian at bond CPs with binding energy.

Further analysis of bonding interactions can be performed using the ratios of individual components of the Hessian. The average directionality of a bond is defined by the curvature of  $\rho(\mathbf{r})$  at bond CPs [25]

$$\langle \tan(\theta) \rangle = \sqrt{\frac{\lambda_1 + \lambda_2}{2\lambda_3}}. \quad (1.4)$$

Directionality can be used to quantify the shape of the electron density at any CP. Large values of directionality indicate that the charge density is close to a radial distribution around the CP. The charge density is more elliptical if directionality decreases, which is indicative of bonding interactions. Values of  $\theta$  in eqn (1.4) have been used to explain material properties such as Cauchy pressure [26], elasticity [14, 25, 27], and brittleness [28].

In a similar fashion, ellipticity is defined using the two negative values of  $\lambda$  at a bond CP

$$\varepsilon = \frac{\lambda_1}{\lambda_2} - 1 \quad (1.5)$$

where  $|\lambda_1| \geq |\lambda_2|$ . A value of  $\varepsilon$  at a bond CP close to 0 indicates a spherical distribution of density, such as in a single bond. The ellipticity will reach a maximum in nonlinear molecules

as the bonding interaction becomes more like a standard double bond and will then drop back to zero as a triple bond is reached (since a triple bond is cylindrically symmetric around the bond path). For the series ethane, benzene, ethene, and ethyne, the ellipticity values at the C–C bond CPs are around 0, 0.19, 0.30, and 0, respectively.

Popkov and Breza were able to explain the selectivity of mono- vs. bi-alkylation of a chiral Ni(II) complex using the ellipticity values at bond CPs. They found that higher mechanical strain led to increased bond CP ellipticities and therefore favored monoalkylation of the complex [29]. Jenkins *et al.* used the values of ellipticity at bond CPs to study the stability of various water clusters. They found a correlation between high ellipticity (moving away from single bonds) and higher energy, less stable water clusters. They also calculated bond path lengths, number of cage CPs, and used the presence or absence of O–O bond CPs to study various water clusters. They concluded that the most stable clusters were actually those with the least number of hydrogen bonding interactions [30].

Values of charge density, the Laplacian, directionality, and ellipticity at bond critical points in the charge density have been used to classify and understand bonding interactions. Since these values are based on the electron density rather than wave functions they can be determined both computationally and experimentally. Values at bond CPs have often been found to correlate to bond strengths and stability of molecules. In general, QTAIM gives insight into thermodynamic properties: stability of molecules and selectivity of reaction pathways.

### 1.1.2 Bond Paths

Bond paths are defined as the union of two gradient paths in  $\rho(\mathbf{r})$  originating at a bond CP and terminating at neighboring nuclei [31]. Additional chemical bonding information can be gained as we move up in dimensionality from the study of zero-dimensional CPs to one-dimensional bond paths. In highly symmetric systems, bond paths lie directly along the internuclear axis, which is the straight line that can be drawn between two bonded nuclei. Bond paths can deviate from the internuclear axis to varying degrees, however, and this

deviation often correlates to the stability of a bonding interaction. For example, hydrogen bond paths usually have at least a slight curvature to them, with weaker hydrogen bonds having more curved bond paths.

In ring structures, it has been argued that bond paths curving outward from the ring indicate ring strain, while bond paths that curve in towards the center of the ring show stabilization [31]. A simple example of this phenomenon can be seen in the C–C bond paths of cyclopropane in Figure 1.2. The bond paths curve out from the internuclear axes, in line with the strong destabilization in this molecule due to ring strain. In benzene, a resonance stabilized structure, the C–C bond paths curve in towards the central ring point.

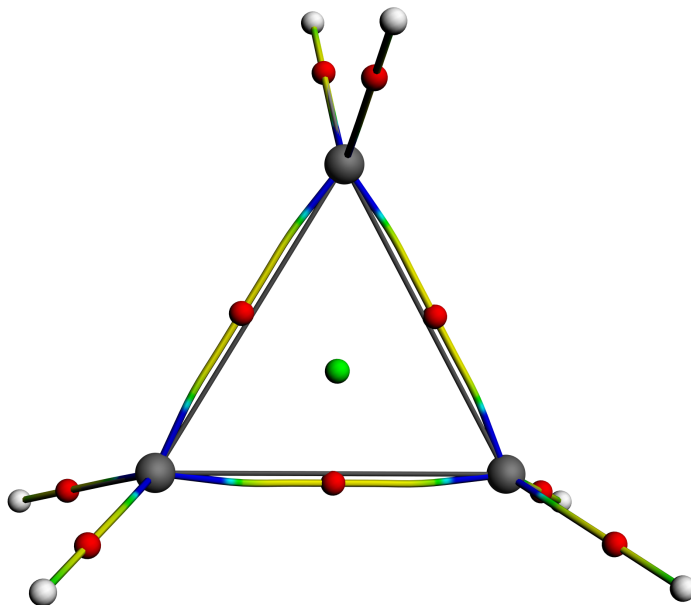


Figure 1.2: Bond paths in cyclopropane. The outward curved bond paths between the carbon atoms (colored lines) indicate ring strain in the molecule. The internuclear axes (grey lines) are shown for comparison.

A bond path is an intuitive way to picture a chemical bond, as a ridge of charge density connecting nuclei. Some chemists have argued that the existence of a bond path is always indicative of a stabilizing bonding interaction [32, 33]. However, there are systems where bond paths are found, yet it has been argued that the interactions between the atoms that

the bond path is connecting are repulsive in nature [17, 34–36]. Additionally, there are instances where a bond path does not exist, yet most chemists would argue that a stabilizing bonding interaction exists [37]. This disagreement over the interpretation of bond paths in  $\rho(\mathbf{r})$  indicates a need to advance the current bonding model.

### 1.1.3 Atomic Basins and The Zero-Flux Surface Condition

While the original QTAIM only defines bonding interactions based on 1D lines, it does define physical boundaries for where 3D “quantum atoms” exist within molecules [1]. An atom in a molecule (AIM), often referred to as a Bader atom or atomic basin, can be realized by picking any point in the charge density and following the gradient path of steepest ascent that will eventually terminate at a nuclear CP. All gradient paths that terminate at a single nuclear CP cover the space that defines an atomic volume. These regions are bounded by surfaces of zero-flux in  $\nabla\rho(\mathbf{r})$ , meaning that no gradient path can cross these surfaces, and the zero-flux surface condition is satisfied,

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0, \text{ for all } \mathbf{r} \in \partial\Omega \quad (1.6)$$

where  $\partial\Omega$  is the boundary of each mononuclear region and  $\mathbf{n}(\mathbf{r})$  is a normal vector. An atom in a molecule can therefore be recognized as the union of a nuclei and its attractive basin in the gradient vector field. This partitioning of space allows additive properties of atoms within molecules to be calculated, such as charge, energy, and volume. These properties can be found by integration over atomic basins.

Atomic basins are classified as quantum atoms because they are defined by a quantum boundary condition in terms of a measurable expectation value, an observable (the charge density) [38]. The fact that these volumes have well-defined properties such as energy is due to the zero-flux surface condition, eqn (1.6). An arbitrary volume in  $\rho(\mathbf{r})$  does not have a well-defined kinetic energy since the local kinetic energy density at any point is ambiguous [39, 40]. The total kinetic energy of a molecule can equivalently be calculated by integration



of the Schrödinger kinetic energy,

$$K(\Omega) = -\frac{\hbar^2}{4m}N \int_{\Omega} dr \int d\tau' [\Psi \nabla^2 \Psi^* + \Psi^* \nabla^2 \Psi] \quad (1.7)$$

or the gradient kinetic energy,

$$G(\Omega) = \frac{\hbar^2}{2m}N \int_{\Omega} dr \int d\tau' \nabla_i \Psi^* \cdot \nabla_i \Psi \quad (1.8)$$

or through any linear combination of these two equations [39].

In order for the kinetic energy to be well-defined,  $K(\Omega)$  must be equal to  $G(\Omega)$ . Locally, eqns (1.7) and (1.8) differ by a term that is proportional to the laplacian,

$$K(\mathbf{r}) = G(\mathbf{r}) - \frac{\hbar^2}{4m} \nabla^2 \rho(\mathbf{r}). \quad (1.9)$$

To compute the difference of the average kinetic energy in a volume,  $\Omega$ , one can integrate eqn (1.9). We use Gauss's theorem to write this difference in terms of a surface integral,

$$K(\Omega) = G(\Omega) - \frac{\hbar^2}{4m}N \int dS(\Omega, \mathbf{r}) \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}). \quad (1.10)$$

The surface integral will be zero for any volume bounded by a zero-flux surface in the gradient of the charge density. This means that any volume in the charge density bounded by a ZFS has a well-defined kinetic energy.

Furthermore, the virial theorem gives the total energy of a system based on the kinetic energy, and can be written in terms of density functional theory (DFT) as [41]

$$E = -T[\rho] - \sum_A X_A \frac{\partial E}{\partial X_A} \quad (1.11)$$

where  $T$  is the total kinetic energy functional and  $X_A$  are the nuclear coordinates. The total energy is also given by

$$E = T[\rho] + V[\rho] \quad (1.12)$$

where  $V$  is the total potential energy functional. At a stationary point (i.e. an equilibrium geometry), the force term goes to zero and the total energy can be calculated as

$$E = -T[\rho]. \quad (1.13)$$

Therefore, in regions over which  $T$  is well-defined, the kinetic energy alone can be used to calculate the total energy of a volume in  $\rho(\mathbf{r})$ .

#### 1.1.4 Delocalization Index

In general, atomic basins are used to study atomic properties. Atomic properties are calculated using one-electron operators and integrating over an atomic basin using

$$O(\Omega) = \langle \hat{O} \rangle_{\Omega} = \frac{N}{2} \int_{\Omega} d\mathbf{r} \int d\tau' [\Psi^* \hat{O} \Psi + (\hat{O} \Psi)^* \Psi] \quad (1.14)$$

where  $\hat{O}$  is any one-electron operator. Common values calculated using eqn (1.14) include atomic populations (used to determine atomic charge, referred to as the Bader charge), energies, electrostatic moments, and atomic volumes.

Of particular interest when discussing bonding in terms of QTAIM is the electron delocalization index (DI) between atomic basins. Atomic basins can be used to calculate a bond order (defined in this case as the number of shared electrons pairs between atoms) by integrating the exchange density over two bonded atomic basins. The DI between two atoms is defined as

$$\delta(A, B) = 2|F^{\alpha}(A, B)| + 2|F^{\beta}(A, B)| \quad (1.15)$$

in which  $F$  is the Fermi correlation defined as

$$F^{\sigma}(A, B) = - \sum_i \sum_j S_{ij}(A) S_{ji}(B) \quad (1.16)$$

where  $S_{ij}(\Omega)$  is the overlap integral between two spin orbitals over an atomic region [22].

Matta *et al.* extended the idea of DI between atoms and calculated the delocalization between two phenyl rings in a biphenyl system to study hydrogen-hydrogen bonding. It was found that the delocalization of electrons between the two rings was actually maximized for a planar configuration rather than a twisted (lower energy) configuration. This is used as part of an argument that the controversial hydrogen-hydrogen bond path in biphenyl is a stabilizing interaction [42].

## 1.2 Topological Models of Kinetic Energy Density

The topology of scalar fields other than  $\rho(\mathbf{r})$  have also been used to create chemical bonding models. Many of these functions strive to define delocalization of electrons and are based on the kinetic energy density. Because these functions focus on kinetic energy, they are often interpreted as providing information on the motion of the electrons. Therefore, delocalization functions generally give insight into how the electrons will rearrange and move, making them useful chemical bonding models for predicting reactivity. Here, I review three of the most prominent delocalization functions: the electron localization function (ELF), the electron localizability indicator (ELI), and the localized orbital locator (LOL) method.

ELF was developed with the goal of defining a more rigorous means of bond classification based on local quantum-mechanical functions that depend on the Pauli exclusion principle. Becke and Edgecombe originally defined the function using spin densities [43], but Savin *et al.* generalized ELF for the total electron density as [44]

$$\text{ELF} = \frac{1}{1 + (\frac{D}{D_0})^2} \quad (1.17)$$

which is mapped onto a finite range of  $0 \leq \text{ELF} \leq 1$ , where

$$\begin{aligned} D &= \tau - \frac{1}{4} \frac{(\nabla \rho)^2}{\rho}, \\ D_0 &= \frac{3}{5} (6\pi^2)^{2/3} \rho^{5/3}, \text{ and} \\ \tau &= \sum_i |\nabla \psi_i|^2. \end{aligned} \quad (1.18)$$

$D$  is the difference between the positive kinetic energy,  $\tau$ , and the von Weizsäcker kinetic energy for non-interacting bosons.  $D_0$  is the kinetic energy of a uniform electron gas. ELF is often interpreted as showing the amount of excess kinetic energy there is due to Pauli repulsion [45, 46].

A topological analysis of ELF can be used to define bonding and non-bonding regions in molecules. As opposed to having nuclear attractors in a topological analysis of the charge density, there are three types of attractors in ELF: core, bonding, and non-bonding, each

with its own attractive basin surrounding an electron pair [43]. Between each attractive basin there are surfaces of zero-flux in  $\nabla ELF$  (see Figure 1.3). Saddle points of the  $(3, -1)$  variety are deemed bifurcation points, and the value of ELF at these points is interpreted as the amount of interaction between adjacent basins, i.e. a measure of delocalization.

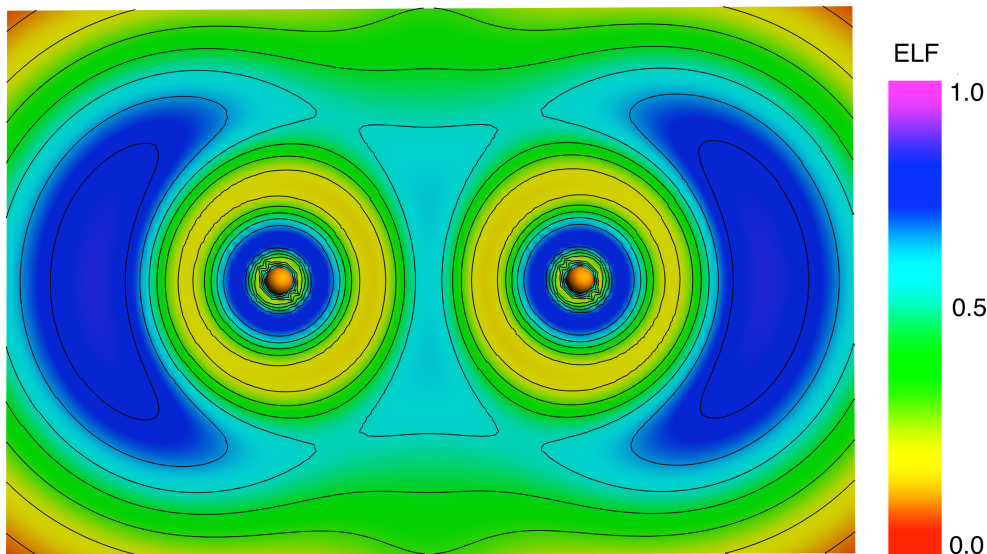


Figure 1.3: A contour plot of ELF for a  $\text{Ge}_2$  molecule showing topological separation of core, valence, and bonding regions into separate regions bounded by ZFSs.

One of the main applications of ELF has been to aromatic compounds. Poater *et al.* recently published a detailed review comparing the use of QTAIM and ELF methods for analyzing aromatic bonding [46]. It has been shown that it is possible to characterize aromaticity in some molecules by separating ELF into  $\pi$  and  $\sigma$  orbital contributions [47]. While the bifurcation values in the total ELF were not able to distinguish between aromatic, anti-aromatic, and non-aromatic systems,  $\pi$  ELF values were. Santos *et al.* used  $\pi$  ELF values to classify a variety of molecules including benzene, a proposed  $\text{B}_6(\text{CO})_6$  molecule, and cyclohexatriene, into each of these categories of aromaticity. Fuster *et al.* performed a study looking at ortho/para versus meta orienting substituents on benzene rings. They were able to distinguish the preferential site of attack (ortho/para or meta) for a second substituent for a set of substituted rings based on isosurfaces of ELF.

It has been argued that ELF is able to predict site reactivity because this function quantifies the importance of Pauli repulsion at a certain point in a molecule. When electrons are unpaired, or in pairs of antiparallel spin, Pauli repulsion is not a large contributor to the total energy. There is not excess kinetic energy in this region compared to a uniform electron gas. At the boundaries between paired electrons though, Pauli repulsion is extremely important causing an increase in the kinetic energy of electrons and low values of ELF.

A recent variation to ELF is the electron localization index D (ELI-D) which analyzes the average number of electrons per fixed fraction of a same-spin electron pair. ELI-D holds the number of same-spin electron pairs fixed for a variable volume cell allowing for different values of charge,  $Q_i$ . Therefore, ELI-D can be interpreted as being proportional to the average number of electrons required to form a fixed fraction of the same-spin electrons. The index has been defined by Kohout and coworkers [48] as

$$\text{ELI-D} = \sum_i^{\text{occ}} \rho_{\sigma,i} \left[ \frac{12}{\rho_{\sigma}(\tau - \frac{1}{4} \frac{(\nabla \rho_{\sigma})^2}{\rho_{\sigma}})} \right]^{3/8} \quad (1.19)$$

where  $\tau$  is the same positive kinetic energy used in ELF (eqn (1.18)), and  $\sigma$  indicates spin.

While this index is topologically identical to ELF, there are some distinctions in the interpretation and calculation. Mainly, ELI-D does not require the use of a uniform electron gas as a reference state. The partial ELI-D (pELI-D) further decomposes this index into individual contributions from each orbital [49]. The pELI-D provides a measure of which orbitals contribute to a portion of  $Q_i$ . Using pELI-D, Grin *et al.* were able to gain additional insight into transition metal bonding, enabling them to quantify the contribution of specific orbitals to bonding in  $\text{Sc}_2^+$  and  $\text{TiF}_4$  [50].

One of the newest topological methods for analyzing bonding interactions is the localized orbital locator (LOL) originally proposed by Becke and Schmider [51]. LOL leaves out the term for the kinetic energy of non-interacting bosons, and only uses the positive form of the kinetic energy [39, 52, 53]. As in ELF, LOL,  $\nu(\mathbf{r})$ , is mapped onto a finite range,  $0 \leq \nu(\mathbf{r}) \leq 1$

using the formula

$$\nu(\mathbf{r}) = \frac{1}{1 + \frac{\tau}{D_0}}. \quad (1.20)$$

LOL is the easiest to compute of the three delocalization formulas presented here, and it has been argued that it is also the easiest to interpret [49].

Rather than focusing on electron delocalization, Jacobsen emphasizes an interpretation of LOL based on the velocity of electrons. A value of  $\nu(\mathbf{r}) = \frac{1}{2}$  means that electrons at  $\mathbf{r}$  are moving at the same speed as they would be if they were a uniform electron gas.  $\nu(\mathbf{r}) < \frac{1}{2}$  indicates high kinetic energy and thus faster electrons. Finally,  $\nu(\mathbf{r}) > \frac{1}{2}$  corresponds to slow moving electrons. Faster electrons are generally found near the nuclei and are thus core electrons. Areas with slower moving electrons indicate valence electrons. LOL was specifically designed to analyze regions of covalent bonding based on the idea that the driving force for covalent bonding is the lowering of kinetic energy [54].

In the original paper introducing the LOL method, covalent bonds and lone pair regions were accurately predicted for simple reactions. This was accomplished by calculating contour plots of  $\nu(\mathbf{r})$  for reactants, transition states, and products [51]. The authors observed that the shape of regions with  $\nu(\mathbf{r}) > \frac{1}{2}$  can be correlated with  $\sigma$  bonding,  $\pi$  bonding, or lone pair regions. It has been shown that LOL generally does a better job of distinguishing between weak and enhanced  $\pi$ -delocalization in aromatic compounds than ELF does, such as in benzene and borazine [49]. Additionally, Yang performed a study using a multitude of bonding metrics to characterize various phthalocyanides. He found that LOL distinguished between core, bonding, and non-bonding regions in a much clearer fashion than ELF [55].

While LOL is still a very new function, there is potential for it to become a powerful tool in predicting reactivity as it focuses on the motion of electrons in molecules. This idea will be developed further when discussing the decomposition of atomic basins into gradient bundles in Chapter 2.

### 1.3 Bond Bundles

Although the only volumes recognized in the original QTAIM are atomic basins, they are not the only meaningful volumes in the charge density bounded by ZFSs and containing well-defined properties. Pendás *et al.* defined a primary bundle as a set of all gradient paths in the charge density that start at one minimum (cage CP) and end at the same maximum (nuclear CP) [13]. Primary bundles are bounded by ZFSs and have one nuclear CP, one cage CP,  $n$  ring CPs, and  $n$  bond CPs on the surface of each bundle. Primary bundles take the shape of polyhedra, with  $2n+2$  vertices and  $2n$  faces. The union of all the primary bundles sharing a single nuclear CP is the atomic basin of attraction discovered by Bader. Pendás *et al.* then went on to define a basin of repulsion as the union of all primary bundles sharing a single cage point.

The elegance of the definitions of atomic and repulsive basins is that there is only one ZFS bounding a nuclear or cage CP that does not pass through the point itself. This makes it relatively easy computationally to locate attractive and repulsive basins in the charge density. Saddle points in  $\rho(\mathbf{r})$  have an infinite number of ZFSs bounding them. A 2-dimensional cut plane of some of the ZFSs surrounding the C–C bond CP in ethene is shown in Figure 1.4.

It was noted by Eberhart that while there are an infinite number of ZFSs bounding bond CPs, some of these ZFSs are special [56] as they are also contained in a 2-dimensional relative critical set of  $\rho(\mathbf{r})$ . Relative critical sets are a higher dimensional extension of critical points applicable to lines and surfaces for a function in  $\mathbb{R}^3$  (see [57, 58]). Relative critical points, which are the elements of a relative critical set, can be defined [57] similarly to the elements of the Eberly height-ridge [59], which is one example of a relative critical set. Let  $I = \{i_1, \dots, i_{3-d}\} \subset \{1, 2, 3\}$ . A point  $\mathbf{r}_0$  is a  $d$ -dimensional relative  $I$ -critical point of  $\rho$  if the eigenvalues of the Hessian,  $H(\rho)$ , are distinct from each other and the eigenvectors,  $e_i$ , for  $i = 1, 2, 3$  of  $H(\rho)$  satisfy the condition

$$\nabla \rho(\mathbf{r}_0) \cdot e_i = 0, \quad \text{for } 1 \leq i \leq 3 - d. \quad (1.21)$$

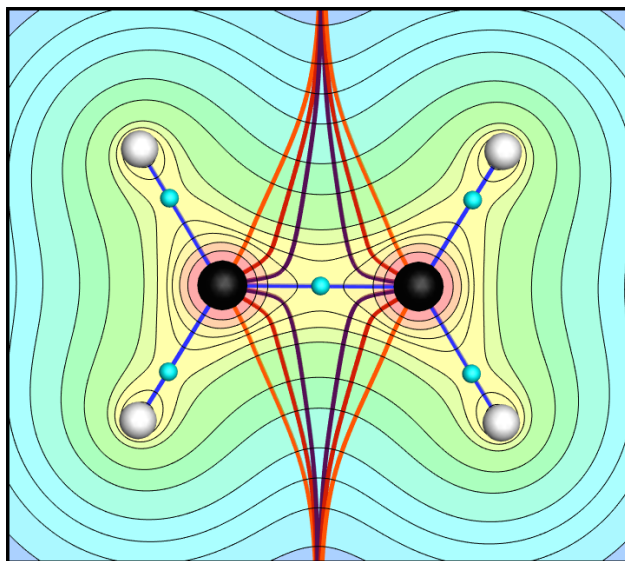


Figure 1.4: 2D cut plane of the charge density of ethene in the plane of the molecule. Thick colored lines show a sampling of ZFSs bounding the C–C bond CP.

The relative  $I$ -critical set of  $\rho$  is the set of all such relative  $I$ -critical points. Note that the number of entries in  $I$ , and thus the number of orthogonality conditions required in eqn (1.21), determines the dimensionality of the critical set. Critical points, lines and surfaces require three, two, and one orthogonality conditions, respectively.

A bond bundle is therefore defined as a volume bounded by ZFSs that satisfy the definition of 2-dimensional relative critical surfaces, and containing a single bond CP. A BB has well-defined, additive properties just like an atomic basin, due to the ZFS condition. Eberhart and Jones have shown that bond bundles recover properties consistent with the standard picture of a bond such as electron count. For example, when the valence charge density is integrated over the bond bundle between bonded carbons for ethane, benzene, ethene, and ethyne; two, three, four, and six electrons are recovered, respectively [60]. This is in agreement with the standard model of these bonds being single, one and a half, double, and triple bonds. Bond bundles have been used to uncover structure–property relationships using two different approaches [61]. The first is to investigate the geometry of ZFSs bounding a BB, and the second is through integration of properties over BB volumes.



As an example of the first approach, Jones showed that the shape and size of bond bundles can be used to predict the tendency of molecules to undergo nucleophilic substitution or addition [62]. Phosphirane (a 3-membered ring) more readily undergoes ring-opening reactions than phosphetane (a 4-membered ring). While the 3-membered ring is more strained than phosphetane, the strain energy alone does not account for the large discrepancy in activation energies for these two reactions [63]. The C–P BB cleaved during the reaction in phosphirane has closed surfaces creating a bond bundle with finite volume. The C–P BB in phosphetane is larger and open on one end. Bond bundles with smaller volumes generally contain fewer valence electrons, and when a bond breaks, a bond bundle must collapse. Therefore, Jones argued that closed BBs will more readily undergo nucleophilic attack, as these bonding interactions are energetically easier to disrupt.

An example of the second approach to the application of bond bundles demonstrates a relationship between electron count and material toughness. The number of valence electrons in BBs was found to correlate with the work of separation,  $W_\infty$ , in a computational study of alloying elements for high strength steels [64].  $W_\infty$  is defined as the minimum amount of energy required to separate an interface into two free surfaces. Finding alloying elements that increase  $W_\infty$  allow for the creation of tougher materials. This study found that the number of valence electrons in second-neighbor Fe–Fe BBs correlated linearly with  $W_\infty$  for an iron-ceramic interface common in dispersion-strengthened steels. This correlation was then used to explain why tougher materials can be created using various transition metals as alloying elements. There is very little  $p_z$  character in the second-neighbor Fe–Fe BBs. By substituting Ni (which possesses low-lying  $p$  character) for some of the Fe sites, the number of electrons in the second-neighbor Fe–Fe BBs is reduced and  $W_\infty$  increases. This prediction is consistent with the known properties of Ni alloyed steels such as BlastAlloy, which is remarkably tough [65]. This is an important example of where a 3-dimensional, physical bond must be studied to understand a property. When the bond CPs between various bonding interactions in the system alone were examined, no correlation was found

with the work of separation.

As these two examples show, a model that presents the chemical bond as both a topological connection between atoms and as a region with well-characterized properties can potentially broaden our understanding of the relationships between chemical structure and properties. The investigation of bond bundles adds a new dimension to ways the charge density can be studied. Rather than looking only at the topology of  $\rho(\mathbf{r})$ , the study of bond bundles involves investigating the geometry of  $\rho(\mathbf{r})$ , specifically by observing the shape of meaningful topological connections. By expanding on this model it may be possible to extract additional useful information from the structure of the charge density. Gradient bundle analysis is being developed with the goal of creating an even higher resolution picture of chemical bonding based on the topology and geometry of  $\rho(\mathbf{r})$ .

## CHAPTER 2

### IN SEARCH OF AN INTRINSIC CHEMICAL BOND

Modified from a paper published in *Computational and Theoretical Chemistry*.

Amanda Morgenstern<sup>1</sup>, Tim Wilson<sup>2</sup>, Jonathan Miorelli<sup>3</sup>, Travis Jones<sup>4</sup>, and M. E. Eberhart<sup>5</sup>. *Comput. Theor. Chem.* 1053: 31–37, 2015.

#### 2.1 Abstract

The chemical bond, as a link between atoms, is an intrinsic property of the charge density. However, bond energy, which is commonly seen as the energy difference between a molecular state and an arbitrary dissociated state, depends extrinsically on the charge density. The view of a bond as a natural link possessing properties that are externally determined often leads to contradictory interpretations as to the origins of the structure and properties of molecules and solids. Ideally, one would like to uncover an intrinsic property of the chemical bond that gives similar information content as that provided by bond energy. To this end, we report on our ongoing work exploring the intrinsic geometry imposed on the charge density by mapping it onto volumes bounded by zero-flux surfaces in the gradient of the charge density, deemed gradient bundles. These natural volume elements of QTAIM have well defined properties. Hence, this mapping produces a set of property distributions with a quantifiable geometric structure that varies from molecule to molecule. Here, we examine the intrinsic geometry of the kinetic energy distribution in gradient bundles for a series of homonuclear diatomic molecules. We find that the structure given by gradient bundles reproduces standard valence shell electron pair repulsion diagrams for the set of molecules,

---

<sup>1</sup>Primary researcher and author

<sup>2</sup>Developed code

<sup>3</sup>Performed distribution statistics and some calculations

<sup>4</sup>Provided advice on computations and document editing

<sup>5</sup>Corresponding author

H<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub>, and F<sub>2</sub>. Furthermore, there is evidence that the geometric properties of these distributions correlate with bond energies.

## 2.2 Introduction

Chemistry is the study of atoms and bonds. QTAIM provides a rigorous and well studied representation of an atom in terms of the intrinsic topology and geometry of the charge density [1]. While QTAIM revealed that a three-dimensional atom with well-defined properties can be unambiguously identified using the topology of the electron charge density, this same theory provides only a one-dimensional representation of the chemical bond, as a bond path. Atomic energies can be calculated by integrating energy values over atomic basins. Since bond paths are only one-dimensional, bond energies can not be obtained in this same manner. Bond energy is generally calculated as the energy difference between the equilibrium conformation of a molecule and some *arbitrary* dissociated state of the same molecule. The bond as a topological linkage is an intrinsic property of the charge density, while its most notable property, the bond energy, is an extrinsic property that depends on the coordinates of the dissociated state. A QTAIM appropriate representation of the chemical bond and its properties must be based in the intrinsic geometry of the charge density. Here we propose one approach toward developing such a representation using gradient bundles in the charge density.

Just as there are charge density ridges (bond paths), there are also one-dimensional critical sets that are formerly termed valley paths [59], which went unnoticed in the original formulation of QTAIM. Because valleys and ridges differ only by the sign of the curvatures along the path, both are often referred to as “ridges”, which is the nomenclature we shall adopt here for any relative critical set. By taking into account all the various types of ridges in the charge density a more complete topological structure of molecules and solids can be constructed in a similar fashion as the primary bundles defined by Pendas *et al.* [13].

For extended systems there will always be four kinds of charge density CPs (0-ridges), six kinds of 1-ridges, and four kinds of 2-ridges [56, 60, 66]. The 1-ridges pairwise connect the

four types of critical points and the 2-ridges are surfaces containing three pairwise connected CPs. This ridge structure forms a set of space filling volumes homeomorphic to tetrahedra. Coincident with the four vertices of each tetrahedron is a nuclear, bond, ring, and cage CP. (Though this is the typical connectivity, tetrahedra can also be formed with a nuclear, cage and two bond or two ring CPs.) The six edges of each tetrahedron are 1-ridges, and the four faces are 2-ridges. In an open system—a molecule or a solid with a free surface—all four types of CPs need not be present, in which case some of the 1-ridges will be of infinite length and some of the 2-ridges will have infinite area [60, 66]. In such systems the 2-ridges form a set of open tetrahedra [61, 62].

In either an open or closed system the tetrahedra are simplices, which means that they may be “glued” together to form a simplicial complex that is homeomorphic to the charge density topology of any molecular system [26, 67]. Accordingly, these simplices have been designated irreducible bundles, IBs, where bundle is used to evoke an image of a bundle of gradient paths (see Figure 2.1). In open systems some of the IBs must be homeomorphic to open tetrahedra and are referred to as open IBs and are said to be open in the direction normal to the tetrahedral face, 2-ridge, of infinite area.

In addition to the extended simplicial complexes, local structures can be generated by gluing together a finite number of IBs. The most basic of these are given through the union of IBs sharing a single CP. The union of all IBs sharing the same nuclear CP generates atomic basins. The union of all IBs sharing the same cage CP yields the repulsive basin first noted by Pendas *et al.* [13]. In addition, one can construct the union of all IBs sharing the same ring point. Finally, there is the union of all IBs sharing the same bond CP. This volume contains a single bond critical point and its associated bond path and recovers the bond bundle volume first identified by Eberhart *et al.* [66]. Figure 2.1 shows the construction of the C–C bond bundle of ethene.

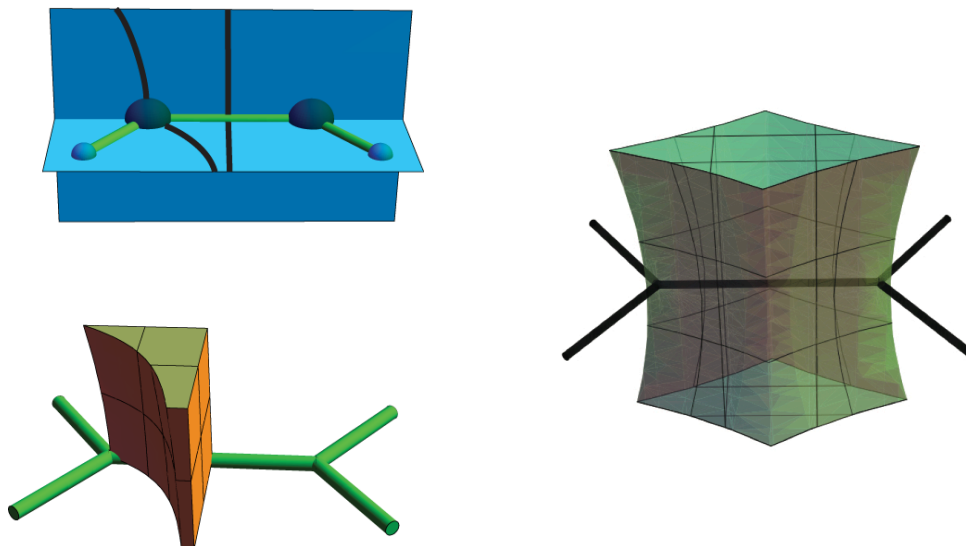


Figure 2.1: A C-C bond bundle in ethene built up from its irreducible bundles. The black lines shown in the top left are the 1-ridges that form the edges of one IB that is a part of the C-C bond bundle in ethene. The 3D IB bounded by ZFSs is shown in the bottom left, and the full bond bundle composed of the eight symmetry equivalent IBs is on the right.

## 2.3 Gradient Bundles

While the IB is the basic structural element capturing the connections between critical points, it too can be built from smaller volumes bounded by ZFSs [67]. We define a gradient bundle as *any* volume in the charge density bounded by a zero-flux surface in  $\nabla\rho(\mathbf{r})$ . To help with visualization, begin by noting that sufficiently close to a nuclear CP all charge density gradient vectors are radial and of the same magnitude. As a result, it is always possible to find a spherical isosurface centered on each nuclear CP.

Figure 2.2 illustrates a method for constructing gradient bundles from a set of neighboring gradient paths. We can triangulate the surface of the spherical isosurface of  $\rho(\mathbf{r})$ . The set of gradient paths passing through the points interior to a triangular element on the sphere form a gradient bundle. Atomic basins, basins of repulsion, and bond bundles can all be decomposed into sets of neighboring gradient bundles, and are in fact gradient bundles themselves. Any property that can be found by integration over one of these larger volumes

can equivalently be calculated by integrating the property over any set of gradient bundles within the larger basin or bond bundle. Gradient bundles can be made arbitrarily small to analyze the charge density to varying degrees of locality.

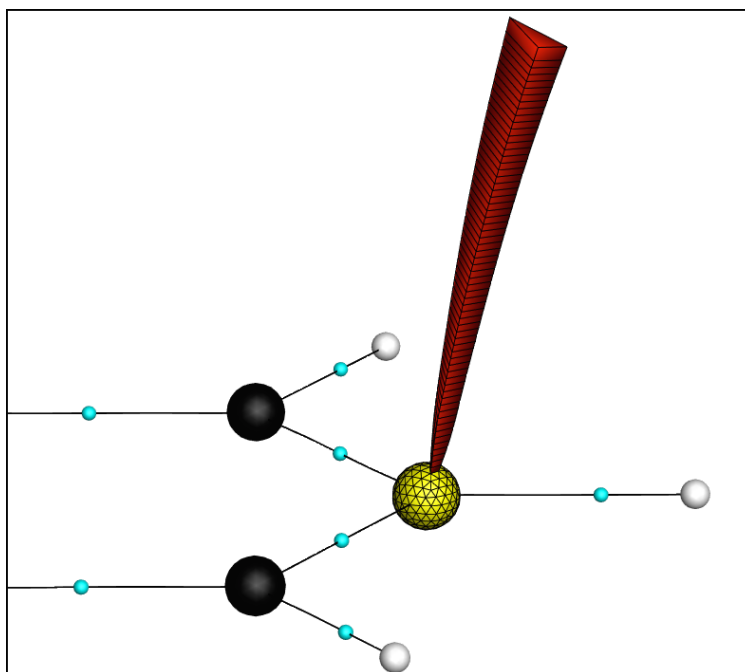


Figure 2.2: A gradient bundle created from the triangulation of a sphere (shown in yellow) around a carbon nucleus in benzene. Carbon nuclei are shown in black, hydrogen nuclei in white, bond critical points in cyan, and bond paths are black.

As motivation for what follows, consider the property distributions of “van der Waals bonded”  $\text{Ne}_2$  and “covalent”  $\text{N}_2$ . Figure 2.3 shows a contour plot of the computed charge densities and the associated gradient fields for each molecule (see Section 2.4 for computational parameters). The gradient fields are visually different, with the  $\text{Ne}_2$  gradient paths being atomic-like (radial) except in the immediate neighborhood of the 2-ridge that is the ZFS of the Ne Bader atoms, i.e., the midplane between the nuclei. In contrast, for  $\text{N}_2$  the gradient path curvature is distributed along a greater length of the path, and particularly, closer to the nucleus. With such distinctive gradient fields, it seems reasonable that their property distributions will also be distinct.

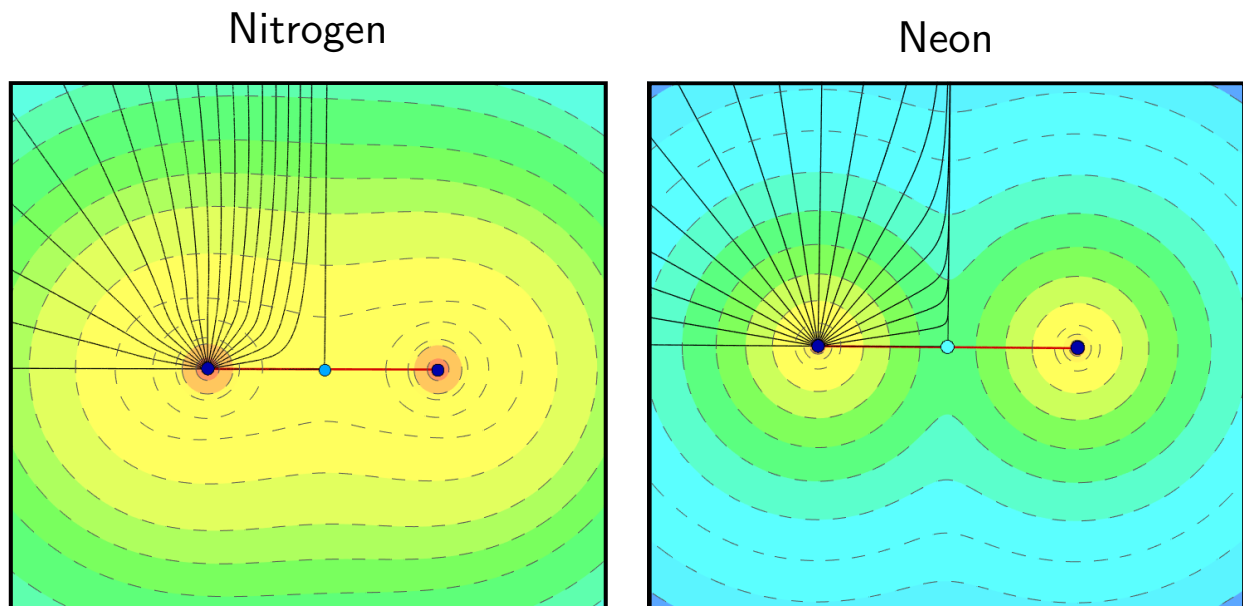


Figure 2.3: Gradient paths seeded every  $10^\circ$  in the charge density for  $\text{N}_2$  and  $\text{Ne}_2$  to show the general gradient path behavior. Dashed lines represent contour lines of  $\rho(\mathbf{r})$ , solid lines are gradient paths, blue circles are nuclear CPs, cyan circles are bond CPs, and red lines are bond paths.

To examine these differences in more detail we explore the intrinsic properties of the chemical bond through a comparison of the charge density and kinetic energy distributions of  $\text{H}_2$ ,  $\text{N}_2$ ,  $\text{O}_2$ ,  $\text{F}_2$ , and  $\text{Ne}_2$ . This series was chosen for the initial study of intrinsic bonding properties based on their symmetry, allowing for easier gradient bundle construction, and their small size, resulting in more expedient calculations.

## 2.4 Methods

The reported results were obtained using the Amsterdam Density Functional Package, ADF, version 2013.01 [20, 68]. All calculations were performed spin-restricted with a non-relativistic all-electron triple  $\zeta$  singly polarized basis set and the Perdew-Burke-Ernzerhof (PBE) parametrization of the generalized gradient approximation (GGA) functional [69–71]. Bond distances for  $\text{H}_2$ ,  $\text{N}_2$ ,  $\text{O}_2$ , and  $\text{F}_2$  were fully optimized using ADF and an internuclear distance of 2.98 Å was used for  $\text{Ne}_2$ . Bader atomic properties were calculated using a grid



based approach implemented in ADF using default parameters which includes using the exact density for the  $V_{xc}$  [72].

In an effort to ensure that our results were not highly dependent on the choice of functional, we also investigated the local density approximation (LDA), Becke-3-Lee-Yang-Parr with 20% hartree-fock exchange (B3LYP), Meta-GGA-type Minnesota 06 local (M06L), and Becke-Lee-Yang-Parr (BLYP) functionals [73–76]. Our principal criteria for functional dependence was visual inspection of the gradient fields and qualitative comparisons of property distributions. For the purposes of comparing properties integrated over gradient bundles, our preliminary data shows that functional choice does not appear to greatly affect the results for this set of molecules (see Appendix A).

We have chosen to investigate only the singlet  $O_2$  molecule to avoid the additional complexity associated with the spin polarized charge density, though it can be treated with QTAIM [77]. We have enforced a “closed-shell” configuration by performing a spin-restricted calculation to approximate the excited singlet  $O_2$  state [78].

The geometry of the charge density was further analyzed using the Bondalyzer add-on package in Tecplot [79]. For each molecule, 73 gradient paths were seeded every  $2.5^\circ$  around a semi circle of radius  $0.05 \text{ \AA}$  centered on a nuclear CP and lying on a plane through the molecule (Figure 2.4). Due to the  $D_{\infty h}$  symmetry of the molecules, the rotation of these gradient paths through any angle about their internuclear axis (by default we used a rotation angle of  $360^\circ$ ) will create ZFSs that bound gradient bundles as shown in Figure 2.5.

The (intensive) properties of rotational gradient bundles are a function only of the azimuthal angle,  $\theta$ , effectively, a 1D property distribution. Here we define  $\theta$  as the angle between each gradient path at its initial seed point and the bond path. Property distributions were found by integrating the calculated values of each property over gradient bundles using a second-order trapezoidal method [79] and a mesh size of  $0.006 \text{ \AA}$  for  $N_2$ ,  $O_2$ , and  $F_2$ ,  $0.005 \text{ \AA}$  for  $H_2$ , and  $0.004 \text{ \AA}$  for  $Ne_2$ . Since the behavior of the gradient paths in hydrogen and neon vary significantly from the other dimers tested (see Section 2.5), a finer mesh was

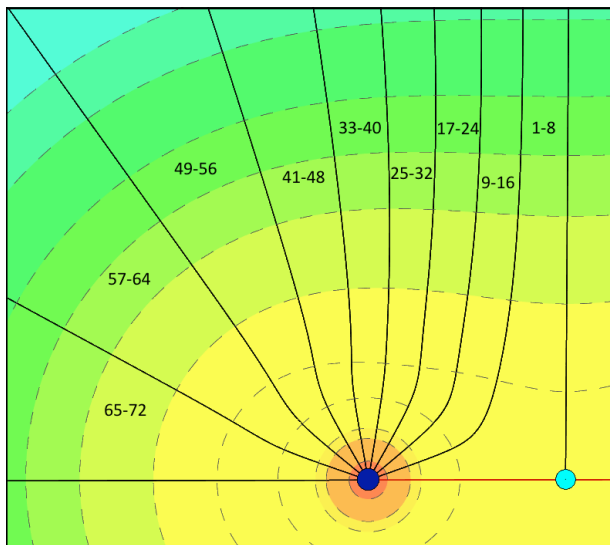


Figure 2.4: In all plots the gradient bundles are numbered in a counterclockwise manner with the first GB containing the bond path and the 72<sup>nd</sup> GB containing the ridge from the nuclear CP in the direction opposite the bond path. Shown for clarity are gradient paths used to create ZFSs every 20°.

used for these molecules after an initial testing at 0.006 Å to ensure the results were not caused by integration errors near the nucleus.

The integration limits extended from 0.05 Å from the nuclear CP to the charge density isosurface of 0.001 *electrons/bohr*<sup>3</sup> for each dimer. In open systems, bader atomic volumes are generally cut off at 0.001 a.u. as this encloses more than 99% of the electron density [1]. To avoid numerical integration errors near the nuclei, the properties from the spherical region within the 0.05 Å radius around the nuclear CP were found separately and proportioned based on volume among the gradient bundles. This is an adequate method of analysis due to the radial behavior of the gradient paths near the nuclei. With this procedure, the integrals were well converged, with electron count and kinetic energy to within 1% of the ADF numerical integration results.

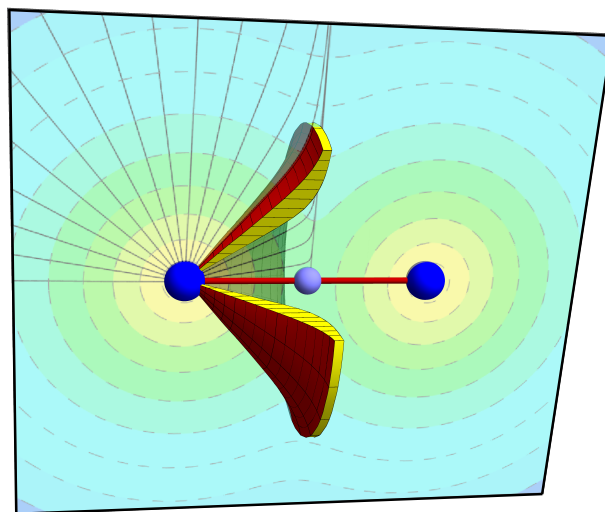


Figure 2.5: A representation of a gradient bundle in  $\text{Ne}_2$  bounded by ZFSs at  $\theta = 40^\circ$  and  $\theta = 50^\circ$  that has been rotated around the internuclear axis by  $280^\circ$ . The image has been tilted to emphasize the structure of a gradient bundle which takes the shape of an umbrella. Note that the gradient bundles in this system will actually extend out to infinity; the GB shown here has been truncated in this image for clarity.

## 2.5 Results and Discussion

The first column of Figure 2.6 plots the integrated electron counts in the rotational gradient bundles for  $\text{H}_2$ ,  $\text{N}_2$ ,  $\text{O}_2$ ,  $\text{F}_2$ , and  $\text{Ne}_2$ .  $\text{N}_2$  and  $\text{H}_2$  are the only molecules with local maxima in the gradient volumes containing the bond path. The other three molecules have greater electron accumulation near the middle of each plot, i.e,  $\theta \approx 90^\circ$ , which is due, in part, to the larger volumes of these gradient bundles.

The most obvious variation through the series is the shift of the mode of the distributions (the maximum) to smaller values of  $\theta$  through the series  $\text{N}_2$ ,  $\text{O}_2$ ,  $\text{F}_2$ , and  $\text{Ne}_2$ . This shift points to an accumulation of electrons in the gradient bundles in the internuclear region, but not necessarily in those along the bond path. Table 2.1 presents the skewness of each plot, which is a useful metric for quantifying the asymmetry of a distribution. Because electron count distributions are essential unimodal ( $\text{N}_2$  is slightly bimodal) the skewness simply indicates on which side of the plot there is a greater number of electrons. Negative

values indicate that the distribution is weighted toward smaller values of  $\theta$ , and positive values toward larger values of  $\theta$ .

Each rotational gradient bundle is bounded by a ZFS and thus satisfies the virial theorem [1]. Accordingly, at the equilibrium internuclear distance  $\langle E \rangle = -\langle T \rangle$  and  $-\frac{\langle V \rangle}{\langle T \rangle} = 2$ , where  $\langle E \rangle$ ,  $\langle T \rangle$ , and  $\langle V \rangle$  are the average values over any volume bounded by a ZFS of the total, kinetic, and potential energies, respectively. Within Kohn-Sham DFT, the kinetic energy is separated into a non-interacting and a correlation kinetic energy,  $T_s$  and  $T_c$ , respectively. While the virial relationship is only satisfied for the total kinetic energy, because a molecule’s  $T_c$  is generally a small component of the total kinetic energy it has often been ignored [72].

Table 2.1: Bond energies and distribution statistics for gradient bundles in homonuclear diatomic molecules.

Molecule	TBE. <sup>a</sup>	EBE <sup>b</sup>	Skewness		
			$n_e$	$T_s^c$	$T_s/n_e$
H <sub>2</sub>	6.75	4.51	1.93	1.043	-0.809
N <sub>2</sub>	16.6	9.80	-0.852	-0.632	-0.357
O <sub>2</sub>	8.59	6.14 <sup>d</sup>	-0.0122	-0.488	-0.395
F <sub>2</sub>	3.51	1.64	0.572	-0.293	-0.0377
Ne <sub>2</sub>	0.0052	10 <sup>-7</sup>	-0.549	-0.538	0.262

<sup>a</sup>Theoretical bond energies (TBE) are from the ADF calculations described in Methods using the PBE functional (Theoretical energies regardless of computational method are known to differ substantially from measured values.). <sup>b</sup>Experimental bond energies (EBE) are reported at 0K and determined thermochemically [80]. <sup>c</sup>All energy values are reported in eV. <sup>d</sup>0.98 eV has been added to the ground-state triplet oxygen EBE to obtain the singlet oxygen EBE [81].

We verified that the virial theorem is satisfied for the Bader atoms in each molecule using the standard Bader analysis package in ADF where  $-\frac{\langle V \rangle}{\langle T \rangle} = 2.0319, 2.0061, 2.0033, 2.0027$ , and  $2.0022$  for H<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub>, F<sub>2</sub>, and Ne<sub>2</sub>, respectively. Furthermore, the correlation energy ( $T_c$ ) for each dimer was found to be less than 1% of the total kinetic energy. It is then tempting to suggest that the average value of  $T_c$  over a gradient bundle will also represent a small contribution to the total kinetic energy over the same volume. Such an assumption would allow us to infer a fairly accurate value of the total energy of a gradient

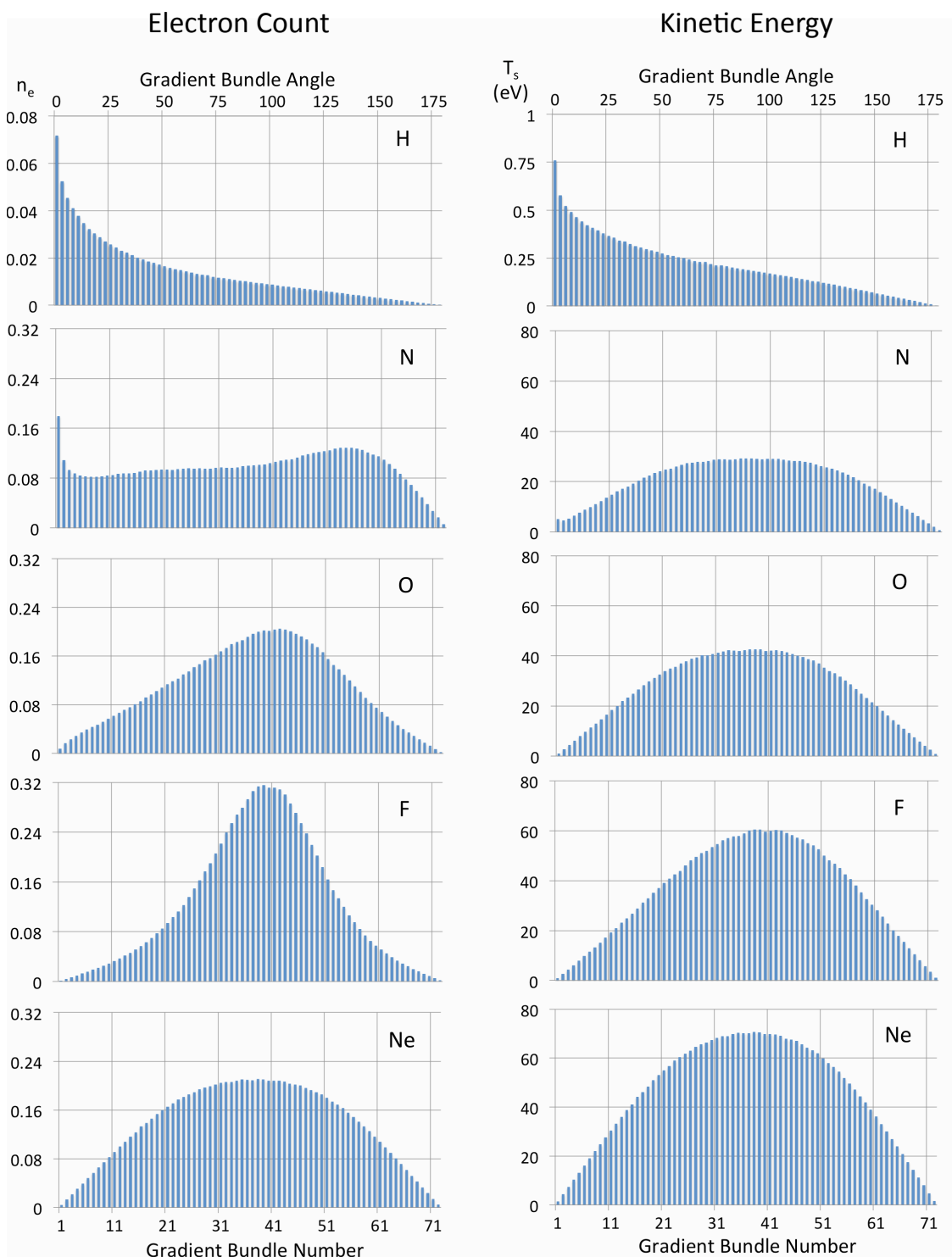


Figure 2.6: Electron and kinetic energy integrations over gradient bundles for a set of homonuclear diatomic molecules. Left: The electron count in each gradient bundle. Right: Kinetic energy density in each gradient bundle in electron volts (eV). Note that the scales on the plots for hydrogen vary from the rest of the dimers.

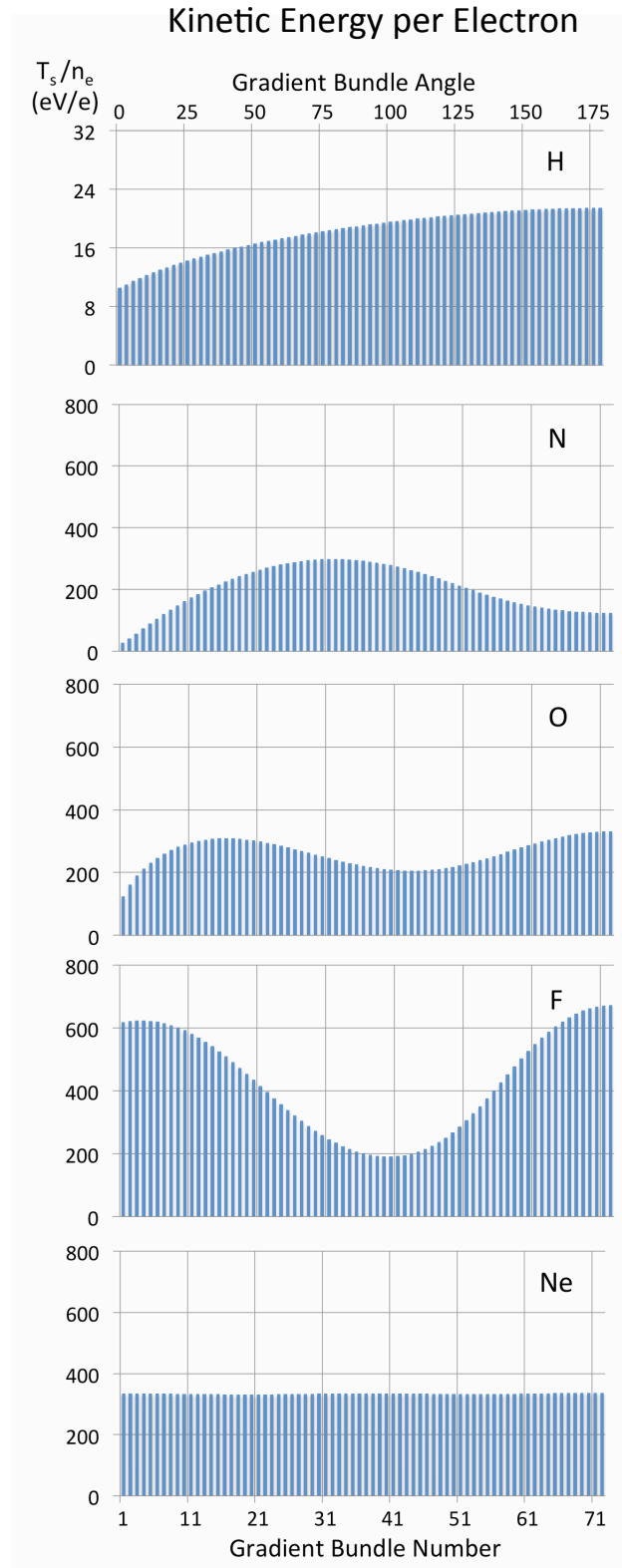


Figure 2.7: Average kinetic energy per electron in each gradient bundle. Note that the scales on the plot for hydrogen vary from the rest of the dimers.

bundle from only a knowledge of the easily calculated non-interacting kinetic energy density,  $T_s$ . However, while  $T_c$  is a small contribution to the total kinetic energy, it is not evenly distributed throughout a molecule due to the freedom in the choice of local kinetic energy [40, 53, 72, 82]. In fact, atomic basins have been shown to sometimes have negative or large contributions of  $T_c$  to the total kinetic energy [72]. It has also been shown that the bond path, and consequently the gradient bundle around the bond path, is a “privileged channel” for exchange and correlation [83]. Therefore, it is likely that the contributions of  $T_c$  to  $\langle T \rangle$  are much greater in some gradient bundles than others. Nonetheless, we have calculated the distributions of  $T_s$  for the five dimers and the results are presented in the right column of Figure 2.6.

Since  $T_s$  is much greater for core than valence electrons, the distribution for  $\text{H}_2$  is significantly different than that of the other dimers.  $\text{H}_2$  lacks any core electrons, creating a unique kinetic energy distribution based purely on valence electrons. For the remaining molecules with core and valence electrons, the plots are dominated by the high kinetic energy of radially distributed core electrons. This artifact is especially noticeable for the rotational gradient bundles near  $90^\circ$ , which contain a greater number of core electrons simply because these gradient bundles have larger volumes.

Inspection of Table 2.1 reveals that the kinetic energy density in the  $\text{H}_2$  gradient bundles has a positive skewness while the other dimers are characterized by a negatively skewed distribution. In the case of  $\text{H}_2$ , this is simply a consequence of the fact that the maximum in the kinetic energy density occurs on the bond path. As a result, the tail of the distribution lies entirely at larger angles. In the other dimers the negative skew shows that the kinetic energy distribution is skewed towards the bond path. For the most part however, the subtle distribution of the valence electrons, and hence bonding, are masked against the background of the core electrons as has been seen in previous studies of kinetic energy densities [52].

This situation can be remedied by calculating the average kinetic energy per electron,  $T_s/n_e$ , by simply dividing the integrated  $T_s$  of each gradient bundle by the total electron

count in the same bundle. Since each core electron will be characterized by the same large kinetic energy (compared to valence electrons) [52], the regions where  $T_s/n_e$  is minimal occur in the gradient bundles with the greatest valence density. The values of  $T_s/n_e$  are plotted in Figure 2.7.

Consistent with this interpretation, there are prominent minima for  $\text{N}_2$ ,  $\text{O}_2$ , and  $\text{F}_2$  near  $180^\circ$ ,  $110^\circ$ , and  $100^\circ$ , respectively. These minima mark the locations of the “lone pairs” as predicted by valence shell electron pair repulsion (VSEPR) diagrams. The single lone pair of  $\text{N}_2$  is accommodated as far as possible from the electrons in the gradient bundle nearest the bond path. The two lone pair of  $\text{O}_2$ , however, are forced through correlation effects into the larger volumes of the rotational gradient bundles near  $110^\circ$ . Finally, the three lone pairs of electrons in  $\text{F}_2$  are located in the deep minimum and larger gradient bundles  $100^\circ$  from the bond path.

While the geometry of the  $T_s/n_e$  manifold at larger angles provides information about the “non-bonding” electrons, it is the geometry of this manifold near the bond path bundles that may correlate with bond energies (see Table 2.1). The most obvious correlation is that bound dimers are characterized by a minimum at  $\theta = 0^\circ$ . The three molecules with the greatest bond energies ( $\text{N}_2$ ,  $\text{O}_2$ ,  $\text{H}_2$ ) have global minima in the gradient bundles containing the bond path.  $\text{F}_2$  possesses a shallow local minimum in the first gradient bundle and this bond is substantially less energetically stable than the other bound dimers. The  $T_s/n_e$  distribution across the gradient bundles in  $\text{Ne}_2$  appears flat, indicating an extremely weak bonding interaction.

## 2.6 Conclusions

While much work remains to be done, our initial results clearly suggest that it is possible to characterize bonding interactions in terms of intrinsic properties of the charge density, and in a fashion consistent with the principles of QTAIM. Central to this approach is the decomposition of a molecular system into gradient bundles. Gradient bundles are the natural volume elements of QTAIM, as their properties are well defined and additive. Using this



approach, we were able to account for variations in the charge density and kinetic energy distribution that are typically assigned to less rigorous decompositions of the charge density, such as its separation into core and valence electrons and the further separation into lone pairs and bonding electrons. This study opens the door to the decomposition of more complex molecules into gradient bundles and the determination of the property distributions of these systems. We fully expect that these distributions will provide a much needed intrinsic representation of the chemical bond.

## CHAPTER 3

### BOND DISSOCIATION ENERGIES FROM GRADIENT BUNDLE ANALYSIS

Modified from a paper published in *Physica Scripta*.

Amanda Morgenstern<sup>1</sup> and M. E. Eberhart<sup>2</sup>. *Phys. Scr.* 91: 023012, 2016.

#### 3.1 Abstract

New and more robust models of chemical bonding are necessary to further our understanding of chemical phenomena. Among these are gradient bundle methods, which analyze bonding interactions in terms of property distributions over geometrically defined volumes in the charge density. Gradient bundle analysis provides a systematic framework from which to search for structure–property relationships. Specifically, the kinetic energy density in gradient bundles can be used to recover the valence electron structure of molecules, and has been demonstrated to qualitatively recover bonding properties in small molecules. Here we present new findings that quantitatively relate the lowering of kinetic energy in bonding regions to bond energies for a set of diatomic molecules. We find a linear correlation between experimental bond dissociation energies and the number of valence electrons in bonding regions defined by kinetic energy density in gradient bundles

#### 3.2 Introduction

Chemical bonding has traditionally been explained by a lowering of potential energy ( $V$ ) as atoms approach their equilibrium distance. However, many chemists argue that covalent bonding is actually due to a lowering of kinetic energy ( $T$ ). This concept may at first appear to be impossible, since the only negative contribution to the energy of a stationary state of a bonded molecule in the Schrödinger equation is due to the attraction between the nucleus and electrons ( $V$ ). However, as atoms go from a non-bonded to a bonded state,

---

<sup>1</sup>Primary researcher and author

<sup>2</sup>Corresponding author

they will not be in a stationary state at all times [54]. In order to better understand the mechanism of chemical bonding one must look at more than the observable states given by eigenvalues from the Hamiltonian at the beginning and end of chemical bond formation. Chemical bonding has been explained by a lowering of  $T$  using three different approaches: 1) observing a reaction coordinate using the virial theorem, 2) energy decomposition methods, and 3) regional distributions of kinetic energy density.

The first of these approaches is due to Slater, who proposed that covalent bonding was due to a lowering of kinetic energy in his early work on the virial theorem [84]. When a molecule is not in an equilibrium state there is a force term on the nuclei, and the virial theorem can be written by taking the force term into account as an addition to the  $T$  and  $V$  terms, or as part of the kinetic and potential energy. Using the latter method one gets

$$\begin{aligned} T &= -E - r \left( \frac{dE}{dr} \right), \text{ and} \\ V &= 2E + r \left( \frac{dE}{dr} \right). \end{aligned} \tag{3.1}$$

Figure 3.1 shows that close to the equilibrium distance,  $r_{eq}$ , there is a decrease in  $V$  as two non-bonded atoms approach one another to form a diatomic molecule. Further from this point along the reaction coordinate (when the two atoms are far away) there is actually a large decrease in  $T$  and an increase in  $V$ , before  $V$  decreases near  $r_{eq}$ . This is due to the large force term in eqn 3.1 as  $r$  is increased.

Klaus Ruedenberg has also long been a proponent that covalent bonding is caused by a lowering of kinetic energy. He makes this argument based on energy decomposition methods [54, 85]. The energy change when two atoms or fragments approach to form a molecule can be decomposed into promotional, interference, sharing penetration, and quasi-classical interactions. The interference energy due to  $T$  is greatly negative. While the total  $V$  decreases and the total  $T$  increases upon bond formation, Ruedenberg argues that the lowering of the interference kinetic energy is what determines whether a covalent bond will form between two atoms as well as determining the total energy change of the system.

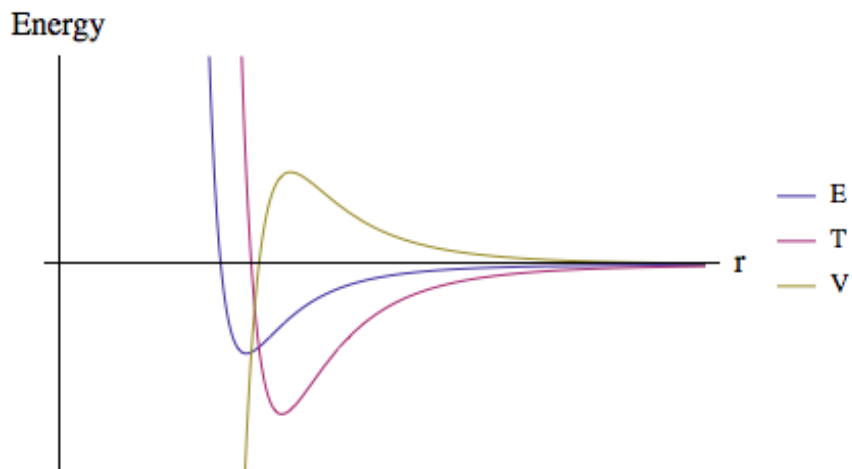


Figure 3.1: A plot of the relationship between total energy, kinetic energy, and potential energy in a homonuclear diatomic molecule using the definitions from eqn 3.1. The total energy was calculated using a pair potential of the form  $-a/r^5 + b/r^9$ .

Finally, functions such as ELF and LOL define bonding interactions based on a lowering of  $T$  in spatial regions [44, 45, 51, 52]. Mapping the kinetic energy density onto a normalized field (see Section 1.2) shows that bonding regions between nuclei and lone pair regions have locally depleted  $T$ . While these methods define volumes corresponding to “bonding regions,” electron integrations in these volumes is rarely performed, as it is difficult to interpret any conceptual meaning tied to these volumes or electron counts. This is because volumes based on the topology of ELF and LOL do not have well-defined energies that can be calculated. ELF and LOL bonding basins are not bounded by ZFSs in  $\nabla\rho(\mathbf{r})$ , but by ZFSs in the gradient of functions based on the kinetic energy density.

In Chapter 2, we found a qualitative relationship between spatial regions with depleted kinetic energy and experimental BDE. Here we present a quantitative relationship between  $T$  and BDE, that is based on well-defined energies in regions bounded by ZFSs in  $\nabla\rho(\mathbf{r})$ . We determine the volumes of gradient bundles in regions with a lowered kinetic energy density for a variety of diatomic molecules. Then, we present a correlation between the number of valence electrons in these regions and experimentally determined BDEs. Gradient bundles thus provide an avenue to further coalesce the ideas of a chemical bond as an energetically

stabilizing interaction and a physical topological connection in  $\rho(\mathbf{r})$ .

### 3.3 Methods

Charge densities and energies were obtained using the Amsterdam Density Functional Package, ADF, version 2013.01 [20, 68]. All molecules were optimized using spin-unpolarized calculations with a non-relativistic all-electron triple  $\zeta$  singly polarized basis set, no frozen core, and the Perdew-Burke-Ernzerhof (PBE) parametrization of the generalized gradient approximation (GGA) functional [69–71].

Values of  $\rho(\vec{\mathbf{r}})$  and  $T_s$  were imported into Tecplot [79] on a grid with 0.015 Å spacing for gradient bundle analysis. All gradient paths were seeded on the surface of a nuclear-centered sphere such that  $75 \pm 5\%$  of the core electrons were contained within the sphere. This ensures a radial distribution of electrons at the seed point while maintaining enough distance from the nuclei that the nuclear cusp does not disturb the gradient path trajectory. Kinetic energies and electron counts inside of this sphere were distributed amongst the rotational gradient bundles as described in Chapter 2.

### 3.4 Results and Discussion

With an infinite number of ZFSs in  $\nabla\rho(\vec{\mathbf{r}})$ , there are an infinite number of ways to partition a molecule into regions with well-defined energies. Gradient bundle analysis takes advantage of this fact by creating property distributions within larger atomic basins or bond bundles. Additionally, GBA requires as inputs only the charge density and values that can, in principle, be calculated as a functional of the charge density, e.g., the non-interacting kinetic energy density [86]. This allows for the determination of chemical properties either theoretically or by using high resolution data from x-ray diffraction experiments.

It has been shown [87] that valence shell electron pair repulsion (VSEPR) diagrams can be recovered by integrating the number of electrons ( $n_e$ ) and the non-interacting kinetic energy ( $T_s$ ) within gradient bundles. Minima in the average kinetic energy per electron (KE/e) correspond to bonding and lone pair regions containing valence electrons. Figure 3.2 shows

the KE/e distribution for carbon monoxide. Gradient paths were seeded in a plane every  $2.5^\circ$  around the nuclei and rotated around the internuclear axis to create rotational gradient bundles.  $T_s$  and  $n_e$  are integrated within each GB producing an atomic basin property distribution. Minima in KE/e occur at  $0^\circ$  for both the carbon and oxygen atoms, which corresponds to the rotational gradient bundles bounded by the interatomic surface (IAS) and the first ZFS moving away from the bond path. A second minimum occurs around each atom at  $180^\circ$ , which corresponds to the rotational gradient bundle in the direction opposite the bond path. These minima denote the bonding and lone pair regions in CO, respectively. Carbon monoxide is generally described as having a triple bond and single lone pair on both the carbon and oxygen atoms, consistent with the GBA picture.

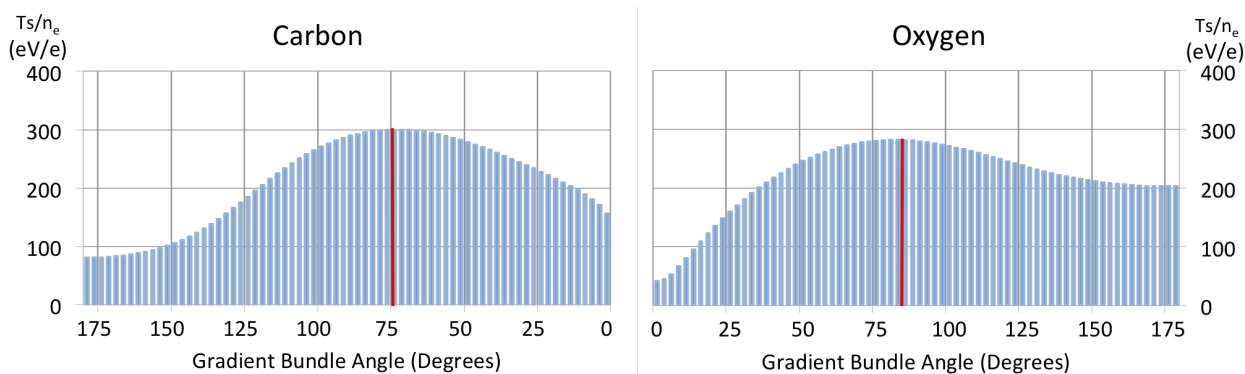


Figure 3.2: Integration of the average KE/e in each  $2.5^\circ$  gradient bundle in CO. The plots are oriented so that they line up geometrically with Figure 3.3: the bonding region is in the center of the image ( $0^\circ$ ) and the lone pair regions are on the outside edges of the plots ( $180^\circ$ ).

It was noted [87] that period 2 homonuclear diatomic molecules with higher bond dissociation energy have a deeper KE/e minimum in the bonding region. Specifically,  $H_2$ ,  $N_2$ , and  $O_2$  have pronounced minima in their KE/e plots at  $0^\circ$ , and are characterized as strongly bound. In contrast,  $F_2$  has a shallow KE/e minimum in the bonding region and a much lower BDE. These results prompted the search for a meaningful relationship between the bonding region KE/e minima in gradient bundles and BDE, a property that has not previously been correlated with the molecular charge density.

As in the LOL and ELF methods, maxima in gradient bundle KE/e plots can be used to define boundaries between bonding and lone pair regions. Kinetic energy bonding regions (KE BRs) are volumes in the charge density bounded by zero-flux surfaces. KE BRs are found by searching for the gradient bundle around each nuclei with the highest average kinetic energy per electron, then using the ZFS bounding this gradient bundle to define the bonding region. The KE BR for CO is highlighted in Figure 3.3.

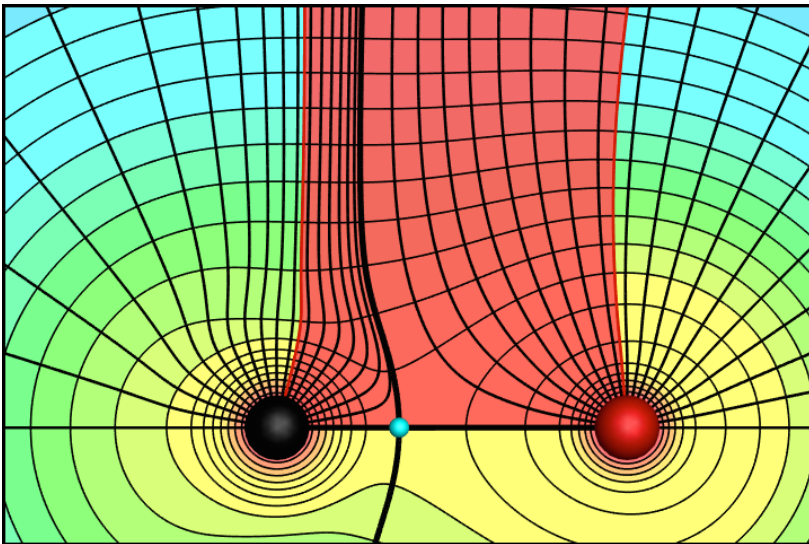


Figure 3.3: A cut plane of the kinetic energy bonding region in carbon monoxide. Contour lines of  $\rho(\mathbf{r})$  are drawn on a logarithmic scale from  $10^{-3} - 10$  e/bohr<sup>3</sup>. Thick black lines show bond paths and interatomic surfaces. Thin black lines show gradient paths seeded every  $10^\circ$  around each nuclei. The red shaded region is a 2D cut plane of half of the KE BR.

Table 3.1 shows the total number of electrons, number of valence electrons, and “normalized” electron counts in KE BRs for a set of diatomic molecules. In the spirit of using only the total charge density (and values directly calculable from  $\rho(\vec{r})$ ) we have used the following method to approximate the number of valence electrons in each bonding region.

First, the total charge density is integrated over the KE BR to obtain the number of electrons in this region. Next, the angle between the bond path and ZFS bounding the KE BR is used to define a spherical sector around the nucleus. Since core electrons maintain a nearly spherical distribution around nuclei, the percentage of a sphere that the spherical

Table 3.1: Total number of electrons, number of valence electrons, and normalized electron counts in kinetic energy bonding regions in a set of diatomic molecules.

Molecule	Total e	Valence e	Normalized e
HF	1.31	1.11	0.56
HCl	3.01	1.70	0.17
HBr	4.06	2.43	0.14
HI	4.57	2.16	0.06
N <sub>2</sub>	6.00	4.43	1.11
O <sub>2</sub>	1.59	1.17	0.29
F <sub>2</sub>	0.04	0.03	0.01
CO	6.01	4.35	1.09
S <sub>2</sub>	4.16	1.82	0.09
Cl <sub>2</sub>	1.52	0.58	0.03

sector occupies determines the percentage of core electrons contained in the KE BR. The approximate number of valence electrons in the KE BR is determined by subtracting the calculated core electrons from the total electron count. Finally, the normalized electron count is determined by dividing the number of valence electrons in the KE BR by the total number of core electrons in the molecule.

The relationship between the normalized electron count of the kinetic energy bonding regions and experimentally determined bond dissociation energies is provided in Figure 3.4 and shows a linear trend with an  $R^2$  value of 0.958. This observation suggests that there is a correlation between the number of valence electrons contained in bonding regions—defined by the kinetic energy density—and BDE. Generally, there will be more electrons in a KE BR if the kinetic energy is depleted over a larger region. Therefore, a greater lowering of kinetic energy in the bonding region corresponds to a stronger chemical bonding interaction.

### 3.5 Conclusions

The gradient bundle model of chemical bonding utilizes the zero-flux surface condition to define bonding regions with well-defined energies, similar to atomic basins. Gradient bundle



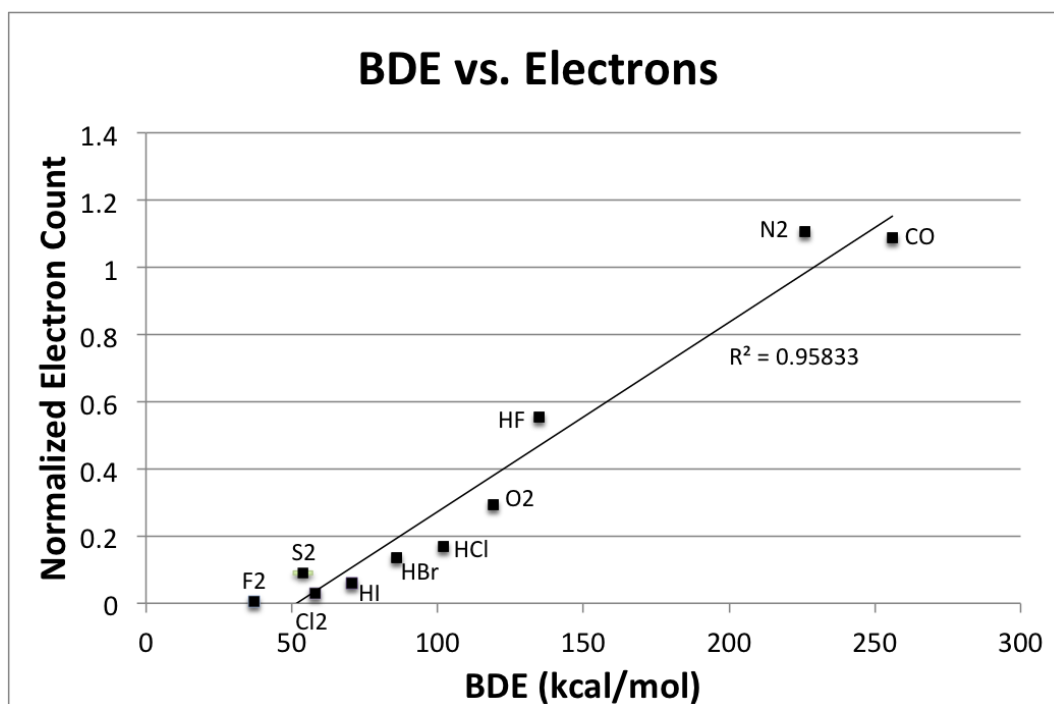


Figure 3.4: Relationship between the normalized electron count in KE BRs and BDE. Experimental BDEs were obtained from [80]

analysis retains the topological picture of a chemical bond given by QTAIM and bond bundle methods. New and existing models should be evaluated based on their ability to rationalize and predict chemical and physical properties. With the extension of gradient bundle analysis, we have uncovered a quantitative relationship between the distribution of kinetic energy in small molecules and their bond dissociation energies, and further rationalized bonding interactions as originating from a lowering of kinetic energy.

By combining the representations of a bond as a stabilizing energetic interaction and as a topological connection in the charge density, we hope GBA will continue to provide useful chemical insights, which will be enhanced through further developments to the gradient bundle model. Specifically, we are searching for structure–property relationships involving concepts such as electronegativity, reactivity, and chemical hardness. Further investigation of the BDE relationship to GBs will involve using energy decomposition methods to investigate the lowering of KE in relation the degree of covalency attributed to specific bonding

interactions. Also of interest is whether or not KE BRs correspond to bond bundle volumes. However, bond bundles are computationally difficult to find with our existing algorithms. This limitation is being addressed through our in-house software development program and we hope to soon be able to compute BBs and GBs for complex low symmetry molecules.

## CHAPTER 4

### THE INFLUENCE OF ZERO-FLUX SURFACE MOTION ON CHEMICAL REACTIVITY

Modified from a paper published in *Physical Chemistry Chemical Physics*.

Amanda Morgenstern<sup>1</sup>, Charles Morgenstern<sup>2</sup>, Jonathan Miorelli<sup>3</sup>, Tim Wilson<sup>4</sup>, and M. E. Eberhart<sup>5</sup>. *Phys Chem Chem Phys*. 91: 5638–5646, 2016.

#### 4.1 Abstract

Visualizing and predicting the response of the electron density,  $\rho(\mathbf{r})$ , to an external perturbation provides a portion of the insight necessary to understand chemical reactivity. One strategy used to portray electron response is the electron pushing formalism commonly utilized in organic chemistry, where electrons are pictured as flowing between atoms and bonds. Electron pushing is a powerful tool, but does not give a complete picture of electron response. We propose using the motion of zero-flux surfaces (ZFSs) in the gradient of the charge density,  $\nabla\rho(\mathbf{r})$ , as an adjunct to electron pushing. Here we derive an equation rooted in conceptual density functional theory showing that the movement of ZFSs contributes to energetic changes in a molecule undergoing a chemical reaction. Using a substituted acetylene, 1-iodo-2-fluoroethyne, as an example, we show the importance of both the boundary motion and the change in electron counts within the atomic basins of the quantum theory of atoms in molecules for chemical reactivity. This method can be extended to study the ZFS motion between smaller gradient bundles in  $\rho(\mathbf{r})$  in addition to larger atomic basins. Finally, we show that the behavior of  $\nabla\rho(\mathbf{r})$  within atomic basins contains information about electron response and can be used to predict chemical reactivity.

---

<sup>1</sup>Primary researcher and author

<sup>2</sup>Provided mathematical guidance on derivation

<sup>3</sup>Provided guidance on DFT concepts and usage

<sup>4</sup>Developed code

<sup>5</sup>Corresponding author

## 4.2 Introduction

A central goal of chemistry is to explain and predict chemical reactions. For the better part of the last century our approach toward achieving this goal has been to seek a better understanding of electron response to a changing chemical potential. This endeavor has produced two complementary views of electron response. The first is an orbital perspective, where we picture electrons moving between occupied and unoccupied orbitals as the nuclear coordinates of reacting molecules change. This powerful perspective forms the foundation of frontier orbital theory [88] and finds productive applications across all branches of chemistry, particularly in the study of coordination compounds. The second approach is density based and views electrons as flowing between and within molecules. The prototypical representation for such flow is provided by electron pushing formalisms where electron flow is represented with curved arrows [89]. This representation of electron response provides the foundation for much of organic chemistry.

While the orbital perspective has profited from advances in computational methods, the density-based approaches have not benefited to the same extent. Although it is now a routine matter to calculate molecular charge densities, discerning the charge flow from these calculations is rarely as apparent as the electron pushing formalism portrays.

Researchers have sought to reconcile electron pushing using calculated and experimentally obtained electron densities [62] with the goal of predicting chemical reactivity. As part of this attempt, alternative ways to picture electron density response have been proposed. Rather than as electron flow, Bader presented chemical reactions as resulting from the movement of critical points in  $\rho(\mathbf{r})$  to produce topological catastrophes [1, 90]. Ayers added to this notion with the electron preceding picture (EPP). In this model, electron density around a critical point is assumed to respond more readily in the most compliant directions, where compliance has been shown to often correlate with the magnitude of the eigenvectors of the Hessian of  $\rho(\mathbf{r})$  [91–93]. From this vantage point, electron density responds more like an inhomogeneous substance subject to deformation than as a flowing fluid.

Though the EPP comes closer to capturing the essence of electron response than electron pushing arrows, this approach does not lend itself to visualization, as electron flow, a vector quantity, has been replaced with the rank two compliance tensor. Additionally, this picture is still imperfect, as critical points are but a single and incomplete characteristic of the charge density’s gradient field,  $\nabla\rho(\mathbf{r})$  [34]. Accordingly, here we explore the response of the charge density’s full gradient field to a perturbation. Our hope is that the understanding resulting from this line of investigation will allow for a closer marriage between the density based electron pushing schemes and the fundamental properties of electron response that can now be investigated using both theoretical and experimental techniques.

Our approach is to use gradient bundle analysis, which reduces the gradient field to a set of regions bounded by zero-flux surfaces (ZFSs) and hence having well defined properties—including energies. While all molecules can be partitioned into sets of gradient bundles, for computational simplicity we have chosen to focus on a linear molecule in this study, 1-iodo-2-fluoroethyne (ICCF). Figure 4.1 shows an example of a rotational gradient bundle in a neutral ICCF molecule. We then exploit conceptual DFT formalisms to describe the energetic changes of these regions during a chemical reaction. In this paper we show that the motion of ZFSs (i.e. the gradient field), between and within atomic basins, plays an important role in energetic changes in molecules and understanding chemical reactivity. Our long-term objective is to improve our ability to predict chemical reactivity using the topology and geometry of the charge density. Here we show that valuable insight into reactivity can be gained by viewing the rearrangement of the charge density alone, without performing (often costly) high level energy calculations.

### 4.3 Background

Conceptual density functional theory (CDFT) provides density based definitions for chemical concepts such as chemical potential, electronegativity, and chemical hardness and softness. These definitions begin with the fundamental equation of density functional theory expressing the energy of a chemical system as a functional of the density. The minimum

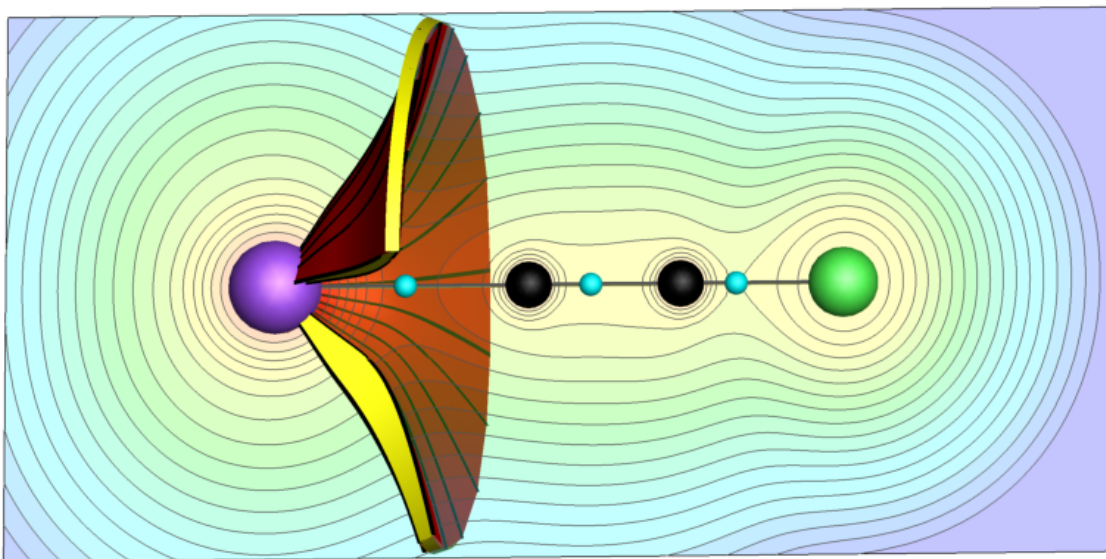


Figure 4.1: A rotational gradient bundle contained in the iodine atomic basin of an ICCF molecule. Contours on a cut plane of  $\rho(\mathbf{r})$  through the molecule are drawn on a logarithmic scale from  $10^{-4} - 10$  e/bohr<sup>3</sup>. Spheres are colored as: I-purple, C-black, F-green, bond CP-cyan; bond paths are grey, and a sampling of gradient paths on the ZFSs are shown in black. This coloring is used in all figures in this chapter.

energy state of a system can be found variationally,

$$\delta(E - \mu\rho(\mathbf{r})) = 0 \quad (4.1)$$

where  $\mu$  is a Lagrange multiplier and has the form of the system’s electronic chemical potential [94], i.e.,

$$\mu = \left( \frac{\partial E}{\partial N} \right)_{\nu(\mathbf{r})} \quad (4.2)$$

where  $E$  is the energy of the molecule,  $N$  is the number of its electrons, and  $\nu(\mathbf{r})$  is the portion of the molecular potential due solely to the atomic nuclei—the so called “external potential.” CDFT associates electronegativity,  $\chi$ , with  $-\mu$ . In general,  $\mu$  is approximated using a finite difference, where one electron is either added or removed from a system. This approach yields two approximations to the chemical potential,  $\mu^+$  and  $\mu^-$ , which find useful applications in the study of nucleophilic and electrophilic attack.

CDFT also recognizes chemical hardness as the linear response of the chemical potential to electron count, which is equivalent to the second derivative of the energy with respect to electron count [95]

$$\eta = \left( \frac{\partial \mu}{\partial N} \right)_{\nu(\mathbf{r})} = \left( \frac{\partial^2 E}{\partial N^2} \right)_{\nu(\mathbf{r})}. \quad (4.3)$$

Eqns (4.2) and (4.3) give global values for chemical potential and hardness. There is no agreed upon definition of atomic analogues of these properties, though several have been proposed [96–102]. A common definition for atomic chemical potential that we will use here is that due to Politzer et al. [97]

$$\mu_i = \left( \frac{\partial E}{\partial N_i} \right)_{N_j, \nu(\mathbf{r})_{i \neq j}} \quad (4.4)$$

where  $N_i$  is the number of electrons associated with atom  $i$ , and  $j$  represents all other atoms in the molecule. This definition can be used in conjunction with a variety of methods that partition charge between atoms—Mulliken populations, Hirshfeld populations, and the populations in the atomic basins of QTAIM can be used to determine  $N_i$ . This definition is also consistent with the electronegativity equalization principle requiring  $\chi$  and  $\mu$  to be constant throughout a molecule at equilibrium [97, 103].

In a similar way we define atomic hardness as

$$\eta_i = \left( \frac{\partial^2 E}{\partial N_i^2} \right)_{N_j, \nu(\mathbf{r})_{i \neq j}} = \left( \frac{\partial \mu}{\partial N_i} \right)_{N_j, \nu(\mathbf{r})_{i \neq j}}. \quad (4.5)$$

During a chemical reaction—when the system’s atoms are no longer in chemical equilibrium and there may be charge transfer—chemical potentials may change. When the charge transfer is small, the new values of atomic chemical potentials are accurately approximated by [104]

$$\mu_i \approx \mu_i^\circ + \eta_i^\circ \Delta N_i \quad (4.6)$$

where  $\mu_i^\circ$  and  $\eta_i^\circ$  are the chemical potential and hardness of atom  $i$  before electron transfer, and  $\Delta N_i$  is the number of electrons transferred to atom  $i$ .

Changes in atomic population, as required by eqn (4.6), can be calculated using condensed Fukui functions. The total Fukui function for a molecule describes the change in electron density resulting from a change in electron count and is given by [105]

$$f(\mathbf{r}, N) = \left( \frac{\partial \rho(\mathbf{r}, N)}{\partial N} \right)_{\nu(\mathbf{r})} \quad (4.7)$$

which can again be approximated with a finite difference, yielding  $f^{+/-}$ . Eqn 4.7 is general, but care must be taken in how the densities are calculated. If degenerate states exist for the  $N$  or  $N \pm 1$  molecule, the Fukui matrix must be calculated using degenerate perturbation theory rather than simply taking an average of the degenerate density functions [106–108]. Integrating the Fukui function over atomic basins defined by QTAIM (or multiplying the Fukui function by atomic weight factors for other atomic population methods, see Ayers et al. [109]) yields a condensed Fukui function and provides the change in atomic electron count.

However, calculating  $\Delta N_i$  in this way does introduce a degree of ambiguity. Specifically, one needs to choose whether to use the response of molecular fragments (RMF) or fragment of molecular response (FMR) approach [109, 110]. For QTAIM atomic basin populations, RMF and FMR methods can give drastically different results [110, 111]. Essentially, the FMR method calculates atomic populations using only the atomic basin from the neutral molecule. For the removal of an electron, the change in electron count for atom  $i$  is calculated by

$$f_i^- = \int_{\Omega_i} \rho(\mathbf{r}, N) dV - \int_{\Omega_i} \rho(\mathbf{r}, N - 1) dV \quad (4.8)$$

where  $\Omega_i$  is the original atomic basin of atom  $i$  in a neutral molecule.

The RMF method uses both the atomic basin from the neutral molecule and the new atomic basin after an electron is removed from the molecule

$$f_i^- = \int_{\Omega_i} \rho(\mathbf{r}, N) dV - \int_{\Omega_i^-} \rho(\mathbf{r}, N - 1) dV \quad (4.9)$$



where  $\Omega_i^-$  is the atomic basin of atom  $i$  in the charged molecule. In the case of QTAIM, this method takes into account the change in the boundary of each atomic basin, i.e., the motion of their bounding ZFSs. The RMF condensed Fukui function gives the difference in the Bader charge of the atom before and after an electron is removed from the system. Advantages and disadvantages of each method have been discussed [110, 112–114]. We believe that the RMF method is more meaningful for studying chemical reactivity, as atomic surface motion plays a role in mediating the energetics of a reaction. In fact, we show that electron response in general is more completely represented by studying the ZFS motion of all gradient bundles as well as those bounding atomic basins.

#### 4.4 Derivation of ZFS Influence on Chemical Reactivity

Here we combine the atomic basin definition provided by QTAIM with CDFT to express the work of atomic surface motion as a contribution to a molecule’s change in energy due to a change in electron count in the system. The motivation for this derivation is in line with our long-term objective, showing that the geometry of the charge density, which can be expressed in terms of the size and shape of its gradient bundles, may be an important tool for understanding chemical reactivity.

While a molecule is undergoing a chemical reaction, the electrons in its atomic basins need not be in an equilibrium state. Electron transfer between atoms is driven by gradients in the chemical potential until a new chemical equilibrium is established with the chemical potential of the atomic basins once again being equal. This could be due to combining two aqueous chemicals, deposition of a gas on a solid surface, or allowing a molecule to interact with an electron reservoir of constant chemical potential. Regardless of the specifics of the process, we can capture its essence by considering an isolated molecule composed of two atoms (1 and 2), which are initially at equilibrium and therefore  $\mu_1 = \mu_2$ , and the ZFS boundary between the two atomic basins is static. We then allow the electron count to vary and determine the change in energy due to this variation. As this energy change is independent of the path, most generally  $\mu_1 = \mu_2$  only at the beginning and end of the charge

transfer. The interatomic distance, and therefore the external potential, is held constant to satisfy the CDFT definition of chemical potential.

For our model diatomic system, energy is a function of the number of electrons associated with atom 1 and atom 2, and the total number of electrons in the molecule is given by  $N = N_1 + N_2$ , such that

$$\Delta N = \Delta N_1 + \Delta N_2. \quad (4.10)$$

During a chemical reaction, the chemical potentials of atoms 1 and 2, which are functions of  $N_1$  and  $N_2$  respectively, will change, as will the number of electrons associated with each atom. The change in energy of the molecule is given by the Taylor series expansion

$$\Delta E = \mu_1 \Delta N_1 + \mu_2 \Delta N_2 + \frac{\partial^2 E}{\partial N_1^2} \frac{\Delta N_1^2}{2} + \frac{\partial^2 E}{\partial N_1 \partial N_2} \Delta N_1 \Delta N_2 + \frac{\partial^2 E}{\partial N_2^2} \frac{\Delta N_2^2}{2} + \dots \quad (4.11)$$

Assuming  $\Delta N_1$  and  $\Delta N_2$  are small, initially we adopt a linear approximation

$$\Delta E \approx \mu_1 \Delta N_1 + \mu_2 \Delta N_2. \quad (4.12)$$

The domain constituting the atomic basin volume of atom 1 is  $\Omega_1$ . Accordingly, the number of electrons in an atomic domain  $\Omega_i$  is given by

$$N_i = \int_{\Omega_i} \rho(\mathbf{r}, N) dV. \quad (4.13)$$

When the number of electrons in the system is changed,  $\rho(r, N)$  will vary as will the boundary separating domains  $\Omega_1$  and  $\Omega_2$ . We denote the new domain of atom 1 as  $\tilde{\Omega}_1$ . Using standard set notation [115], the volume that is common to both the original and new basin is defined as  $\Omega_1 \cap \tilde{\Omega}_1 = \Omega_1^{com}$ . The addition of any volume to the original  $\Omega_1$  is  $\Omega_1^+ = \tilde{\Omega}_1 \setminus \Omega_1^{com}$  and any volume that is removed from the original basin upon changing the number of electrons in the system is  $\Omega_1^- = \Omega_1 \setminus \Omega_1^{com}$ . The new number of electrons in  $\tilde{\Omega}_1$  can be approximated by

$$\begin{aligned}
N_1 + \Delta N_1 &\approx \int_{\tilde{\Omega}_1} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV \\
&= \int_{\Omega_1} \rho(\mathbf{r}) dV + \int_{\Omega_1} \Delta N f(\mathbf{r}) dV + \int_{\Omega_1^+} \rho(\mathbf{r}) dV \\
&\quad + \int_{\Omega_1^+} \Delta N f(\mathbf{r}) dV - \int_{\Omega_1^-} \rho(\mathbf{r}) dV - \int_{\Omega_1^-} \Delta N f(\mathbf{r}) dV
\end{aligned} \tag{4.14}$$

where  $f(\mathbf{r})$  again indicates the Fukui function, and both  $f(\mathbf{r})$  and  $\rho(\mathbf{r})$  are for  $N$  number of electrons for the remainder of the derivation. Eqn (4.14) is exact if the Fukui function is calculated using a finite difference with the same  $\Delta N$  that is used here. The change in the number of electrons in  $\Omega_1$  is

$$\Delta N_1 \approx \int_{\Omega_1} \Delta N f(\mathbf{r}) dV + \int_{\Omega_1^+} \rho(\mathbf{r}) dV + \int_{\Omega_1^+} \Delta N f(\mathbf{r}) dV - \int_{\Omega_1^-} \rho(\mathbf{r}) dV - \int_{\Omega_1^-} \Delta N f(\mathbf{r}) dV. \tag{4.15}$$

Substituting this into eqn (4.12) and combining with eqn (4.10) gives the change of energy in the system as

$$\begin{aligned}
\Delta E &\approx \mu_2 \Delta N + (\mu_1 - \mu_2) \left( \int_{\Omega_1} \Delta N f(\mathbf{r}) dV + \int_{\Omega_1^+} \rho(\mathbf{r}) dV + \int_{\Omega_1^+} \Delta N f(\mathbf{r}) dV \right. \\
&\quad \left. - \int_{\Omega_1^-} \rho(\mathbf{r}) dV - \int_{\Omega_1^-} \Delta N f(\mathbf{r}) dV \right)
\end{aligned} \tag{4.16}$$

where  $\mu_1 - \mu_2$  is the driving force for charge transfer during a reaction. Noting from eqn (4.6) that

$$\mu_1 - \mu_2 \approx \eta_1^\circ \Delta N_1 - \eta_2^\circ \Delta N_2 \tag{4.17}$$

we can combine eqns (4.6), (4.16), and (4.17) to approximate the total change in energy as the sum of three terms

$$\Delta E \approx \Delta E_{\Omega_2} + \Delta E_{\Omega_1} + \Delta E_{ZFS} \tag{4.18}$$

where

$$\begin{aligned}
\Delta E_{\Omega_2} &= (\mu^\circ + \eta_2^\circ \Delta N_2) \Delta N, \\
\Delta E_{\Omega_1} &= (\eta_1^\circ \Delta N_1 - \eta_2^\circ \Delta N_2) \left( \int_{\Omega_1} \Delta N f(\mathbf{r}) dV \right), \text{ and} \\
\Delta E_{ZFS} &= (\eta_1^\circ \Delta N_1 - \eta_2^\circ \Delta N_2) \left( \int_{\Omega_1^+} \rho(\mathbf{r}) dV + \int_{\Omega_1^+} \Delta N f(\mathbf{r}) dV \right. \\
&\quad \left. - \int_{\Omega_1^-} \rho(\mathbf{r}) dV - \int_{\Omega_1^-} \Delta N f(\mathbf{r}) dV \right). \tag{4.19}
\end{aligned}$$

As a particularly convenient way to associate meaning with the above equation, imagine the reaction occurring in three steps. First, electron density from a nucleophile (electrophile) is transferred into (from)  $\Omega_2$ , changing its chemical potential. The energy change associated with this process is  $\Delta E_{\Omega_2}$ . Next, as captured in  $\Delta E_{\Omega_1}$ , the chemical potential difference resulting from the original charge transfer will drive electron flow between  $\Omega_1$  and  $\Omega_2$ . Finally, as part of the charge flow process, the atomic surface between atoms 1 and 2 may move as required by  $\Delta E_{ZFS}$ , which gives the work of boundary motion.

We now rearrange eqn (4.16) by combining terms that depend only on atom 1 and those that depend only on atom 2 giving,

$$\begin{aligned}
\Delta E &\approx \mu_1 \left( \int_{\Omega_1} \Delta N f(\mathbf{r}) dV \right) + \mu_2 \left( \int_{\Omega_2} \Delta N f(\mathbf{r}) dV \right) \\
&\quad + \mu_1 \left( \int_{\Omega_1^+} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV - \int_{\Omega_1^-} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV \right) \\
&\quad + \mu_2 \left( \int_{\Omega_2^+} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV - \int_{\Omega_2^-} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV \right). \tag{4.20}
\end{aligned}$$

The first two terms give the change in energy of the molecule due to the change in the number of electrons in the original atomic basins, modulated by the chemical potentials of atoms 1 and 2, respectively. The last two terms provide the energetic change due to the number of electrons that are in the regions of the atomic basins that have changed due to  $\Delta N$ , modulated by the same chemical potentials. From eqn (4.20), we can extend the linear

approximation of  $\Delta E$  to a molecule with any number of atomic basins bounded by ZFSs as

$$\Delta E \approx \sum_i \mu_i (\Delta N_{\Omega_i} + N_{\Omega_i^{+/-}}) \quad (4.21)$$

where

$$\Delta N_{\Omega_i} = \int_{\Omega_i} \Delta N f(\mathbf{r}) dV \quad (4.22)$$

and

$$N_{\Omega_i^{+/-}} = \int_{\Omega_i^+} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV - \int_{\Omega_i^-} (\rho(\mathbf{r}) + \Delta N f(\mathbf{r})) dV. \quad (4.23)$$

Noting that  $\Delta N_{\Omega_i} + N_{\Omega_i^{+/-}}$  gives the total change in electron count in an atomic basin (using the RMF method), eqn (4.21) can also be written as the sum of atomic energies within a molecule

$$\Delta E = \sum_i \Delta E_i. \quad (4.24)$$

In order to assess the magnitude of the terms in eqn (4.21), we consider the substituted acetylene, 1-iodo-2-fluoroethyne. This molecule was chosen due to the difference in chemical hardness and electronegativity of iodine and fluorine. We hypothesize that the mobility of ZFSs will vary depending on the hardness of the region. Our approach is to use a finite difference to compute the Fukui function and perform the integrations in eqns (4.22) and (4.23) for both  $[\text{ICCF}]^-$  and  $[\text{ICCF}]^+$ .

## 4.5 Methods

All reported charge densities were obtained within the Amsterdam Density Functional Theory package [20, 68] utilizing the M06-2X functional with a QZ4P basis set, no frozen core electrons, and a scalar ZORA approximation to account for relativistic effects [116–119]. The standard Bader analysis package was run on the computed charge densities to obtain critical point locations and atomic charges [120, 121].  $N \pm 1$  calculations were performed unrestricted, with spin = 1, and charge =  $\pm 1$ . Zero-flux surfaces were calculated using the Bondalyzer add-on package in Tecplot [79]. A grid spacing of 0.004 Å was used for

all molecules and electron counts were calculated using the standard integration package in Tecplot following the methods described in Chapter 2.

## 4.6 Results and Discussion

Table 4.1 reports the integration values of the charge density for the terms in eqn (4.21) for  $[\text{ICCF}]^+$  and  $[\text{ICCF}]^-$ . The integrals in eqn (4.23) were determined by finding the volumes between the neutral and  $N \pm 1$  atomic boundaries and integrating  $\rho(\mathbf{r}, N \pm 1)$  over these regions. Eqn (4.22) is simply the FMR condensed Fukui function, giving the change in electron count in the original atomic basin. Figure 4.2 shows the atomic surfaces for  $[\text{ICCF}]^-$  and  $[\text{ICCF}]^+$  compared to the neutral molecule while holding the nuclear coordinates fixed.

Table 4.1: Integration of electrons for the integrals in eqn (4.21) for  $[\text{I-C}_1\text{-C}_2\text{-F}] \pm 1$  electron.

$N - 1$	I	C <sub>1</sub>	C <sub>2</sub>	F
$\Delta N_{\Omega_i}$	-0.576	-0.076	-0.223	-0.108
$N_{\Omega_i^+}$	0	0.018	0.037	0.019
$N_{\Omega_i^-}$	0.016	0.035	0.022	0.002
$\% \Delta N_{ZFS}$	2.7	41.2	20.8	16.5
$\Delta N_i$	-0.592	-0.092	-0.208	-0.090
$N + 1$	I	C <sub>1</sub>	C <sub>2</sub>	F
$\Delta N_{\Omega_i}$	0.678	0.068	0.135	0.073
$N_{\Omega_i^+}$	0.012	0	0.112	0.001
$N_{\Omega_i^-}$	0	0.107	0.001	0.017
$\% \Delta N_{ZFS}$	1.8	61.3	45.7	20.0
$\Delta N_i$	0.691	-0.039	0.246	0.057

The FMR condensed Fukui function does not fully capture the electron response of a molecule. There are at least two other factors that play a role in electron response in the electron density picture. Firstly, contributions to the change in electronic populations in the atomic basins from the surface motion are not negligible. Secondly, electron rearrangement in the original basin can occur without producing a dramatic effect on the integrated electron

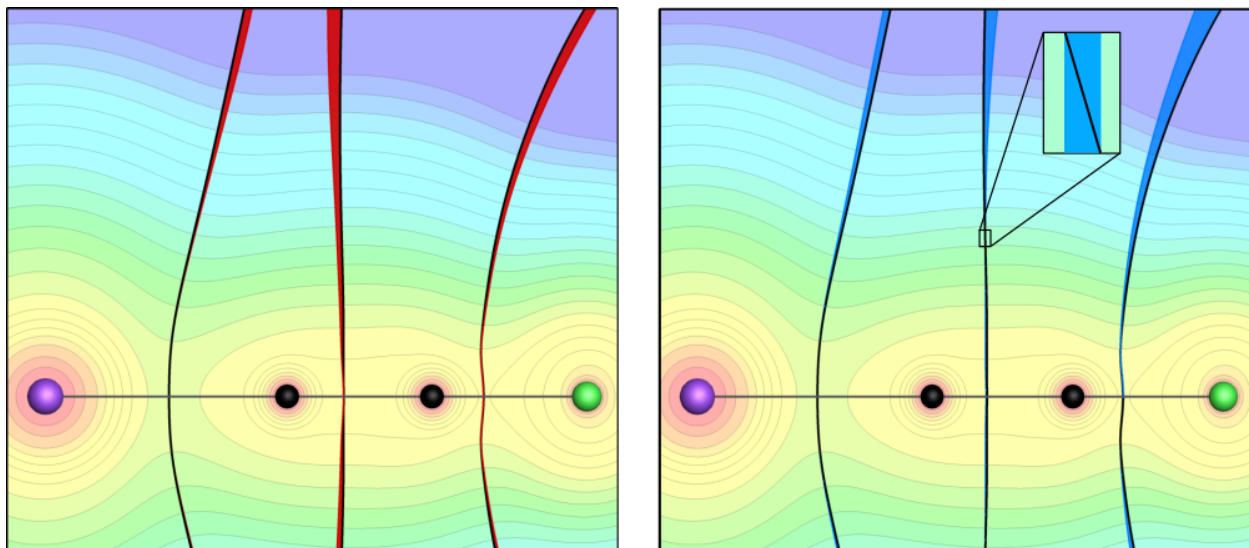


Figure 4.2: Atomic surface motion for ICCF  $\pm 1$  electron. Left: ICCF anion interatomic surface motion. Right: ICCF cation interatomic surface motion. Red (blue) regions indicate the motion of atomic boundaries for the N+1 (N-1) molecule, black lines show the neutral molecule interatomic surfaces. Contours of  $\rho(\mathbf{r})$  on the cut plane are drawn for the neutral molecule.

count within the volume.

We consider the percent contribution of surface motion to the total electron response in the system using the following equation

$$\% \Delta N_{ZFS} = 100 \left( \frac{N_{\Omega_i^+} + N_{\Omega_i^-}}{|\Delta N_{\Omega_i}| + N_{\Omega_i^+} + N_{\Omega_i^-}} \right). \quad (4.25)$$

The numerator provides the number of electrons that are both lost and gained in each atomic basin due to surface motion. The denominator gives the magnitude of the change in electron count within the original basin (the FMR condensed Fukui function) as well as the separate contributions of the atomic basin gaining and losing electrons due to boundary motion. The values of  $\% \Delta N_{ZFS}$  for ICCF range from 1.8 – 61.3%. For  $C_1$  in the  $[ICCF]^-$  molecule, the motion of its ZFS contributes more to this atom's changing electron count than the change in the original basin.

The total change in electron population for atom  $i$  is given by

$$\Delta N_i = \Delta N_{\Omega_i} + N_{\Omega_i^+} - N_{\Omega_i^-} \quad (4.26)$$

which is the RMF condensed Fukui function—the change in Bader charge of the atoms in a molecule resulting from a change in electron count. Substituting the  $\Delta N_i$  values into eqn (4.21) for  $\text{ICCF} \rightarrow [\text{ICCF}]^+$  gives

$$\Delta E_{N-1} \approx \mu_I(-0.592) + \mu_{C_1}(-0.092) + \mu_{C_2}(-0.208) + \mu_F(-0.090) \quad (4.27)$$

and for  $\text{ICCF} \rightarrow [\text{ICCF}]^-$

$$\Delta E_{N+1} \approx \mu_I(0.691) + \mu_{C_1}(-0.039) + \mu_{C_2}(0.246) + \mu_F(0.057). \quad (4.28)$$

As would be expected, the chemically softer halogen, iodine, has a larger change in electron population compared to the harder fluorine atom for the addition or removal of an electron.

It is important to keep in mind that eqns (4.27) and (4.28) are the linear approximations of  $\Delta E$ , and not valid for large values of  $\Delta N_i$ . Our preliminary results suggest a higher order response including chemical hardness and further derivatives would be required to accurately calculate  $\Delta E$ . Incrementally adding fractions of electrons to the system and calculating the change in chemical potential over time, based on atomic hardness, gives an even better approximation of the change in energy. In this initial study however, we are interested in relative electron count changes of regions bounded by ZFSs in a molecule, which would ultimately contribute to the energy change of the molecule based on the chemical potential and higher order terms.

Interestingly, the motion of the  $C_1$ – $C_2$  bond critical point in  $\text{ICCF} \rightarrow [\text{ICCF}]^+$  does not fully represent the atomic surface motion. The inset in Figure 4.2 shows how in some regions the boundary motion between the  $C_1$  and  $C_2$  atomic basin has added volume to  $C_1$ , while in other regions volume has been added to  $C_2$ . The bond critical point between the two carbon atoms, and therefore the atomic surface near the bond path, has moved towards  $C_1$ . But away from the bond path, the atomic ZFS has actually moved into the original  $C_2$  basin. This is further evidence that observing critical point motion alone does not fully represent



the density picture of electron response.

While it is clear that the motion of atomic surfaces plays a role in determining the energetics of a reaction, atomic ZFS motion is also not the only factor in understanding and predicting chemical reactivity. Figure 4.2 shows that atomic surfaces do not always move by a large amount. In fact, the atomic surface of the softer iodine atom moves less than the harder fluorine atom ZFS, especially for the  $N - 1$  case, contrary to what one might expect. For both the addition and removal of an electron, the electron density in  $\Omega^{+/-}$  contributes less than 3% of the total change in electron population for iodine. The movement of interatomic surfaces alone does not correctly depict the electron response.

Significant chemistry is also contained in the rearrangement of electron density within atomic basins. The large  $\Delta N_i$  values for iodine are due to a change in electron count within the original iodine basin. However, electron density does not necessarily add to the original atomic basins in an equally distributed manner. The electron response within atomic basins can be seen more clearly in a cut plane of ZFSs that form gradient bundles for ICCF before and after the change in electron count (Figure 4.3).

For the  $N+1$  system, there is a large amount of electron rearrangement within the original iodine basin. This is not the case for the  $N - 1$  molecule. This result can be rationalized by returning to the frontier orbital view. Adding an electron to ICCF should put electron density into the LUMO while removing an electron should cause electron density to be lost from the HOMO of the neutral molecule. Figure 4.4 shows that the HOMO of ICCF is concentrated most around the iodine atom, then the carbon atoms, and has the smallest contribution near the fluorine atom. Around each nuclei, the density is fairly equally distributed causing little redistribution of the electron density within each atomic basin. Density is removed in a relatively uniform fashion around each nuclei.

The concentration of the LUMO for ICCF, however, varies throughout the molecule. The majority of electron density is added to a portion of the iodine basin, in the region pointing away from the rest of the molecule, as this is where the LUMO is concentrated. The addition

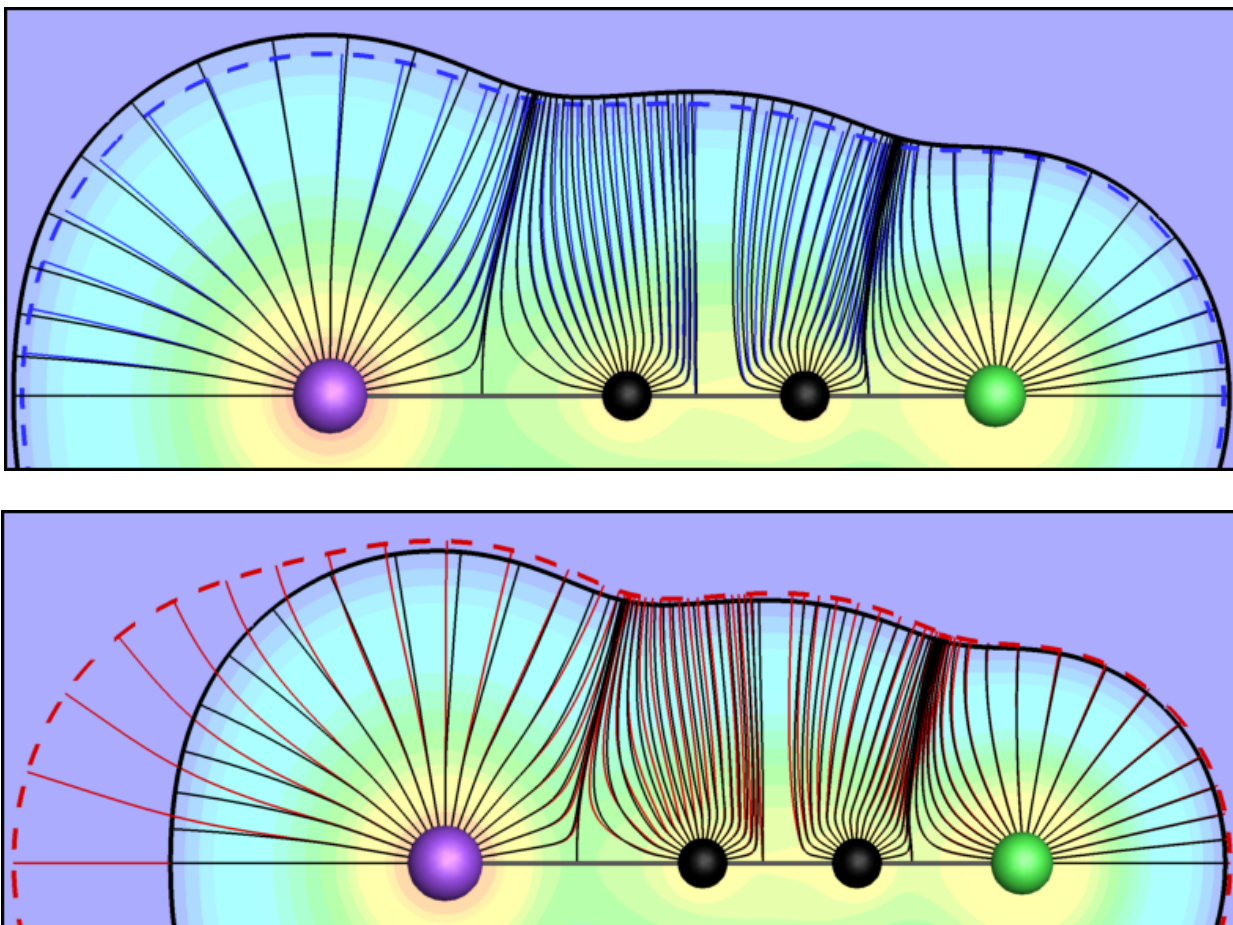


Figure 4.3: Overlays of a cut plane of ZFSs for the top half of a neutral ICCF and  $[\text{ICCF}]^+$  (top) and  $[\text{ICCF}]^-$  (bottom) molecules. Blue (red) lines represent ZFSs for the  $N-1$  ( $N+1$ ) molecule and black ZFSs are for the neutral molecule. Contour lines at  $10^{-3} \text{ e/bohr}^3$  are shown in solid black for the neutral molecule and dashed blue (red) for the cation (anion). Contour flooding of  $\rho(\mathbf{r})$  on the cut plane is for the neutral molecule.

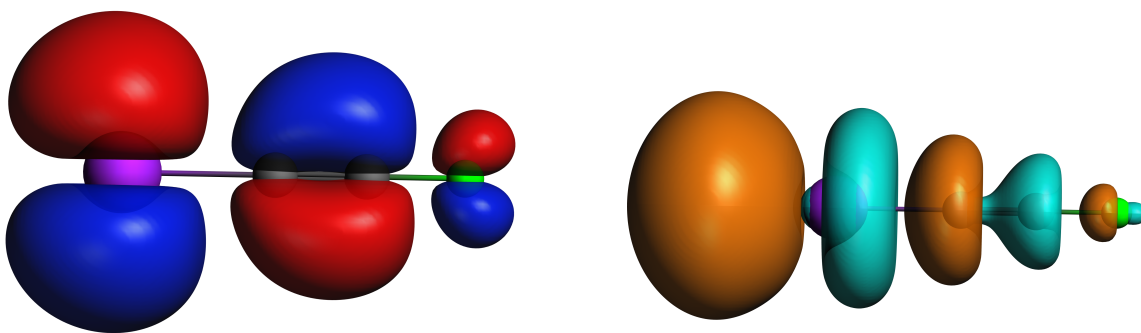


Figure 4.4: The HOMO (left) and LUMO (right) for a neutral ICCF molecule.

of an electron causes a large rearrangement of the electron density within the iodine basin. The electron response can be seen by observing the curvature of the zero-flux surfaces within atomic basins. When  $\text{ICCF} \rightarrow [\text{ICCF}]^-$ , the curvature of the ZFSs in the iodine atom switch from positive to negative in the region pointing away from the I-C<sub>1</sub> bond path.

Thus far, we have shown that eqn (4.21) can be used to represent energetic changes in atomic basins. However, there is no reason to limit energy calculations to atomic regions. This equation can be used for any volume with well-defined energies, i.e., any gradient bundle. In order to gain a highly localized picture of electron response, eqn (4.21) can be used to study the energetic changes in gradient bundles. Preliminary results show that the motion of ZFSs bounding gradient bundles can be used to understand and predict electron response, obtaining similar results to frontier orbital theory.

## 4.7 Conclusions

Electron response can be viewed from an orbital or an electron density perspective. A common way to rationalize and predict chemical reactivity from the density viewpoint is to use electron pushing arrows. The electron pushing formalism is a powerful tool for the prediction of chemical reactions, but it does not give an entirely accurate picture of electron rearrangement. Individual electrons do not flow between atoms and chemical bonds as chemical reactions occur. Therefore, we propose viewing electron response as the motion of zero-flux surfaces rather than electrons flowing between and within molecules.

Using concepts from QTAIM and conceptual DFT, we have shown that the motion of atomic boundaries plays a role in determining energetic changes of molecules when undergoing reactions. This method of rationalizing electron response is in line with the response of molecular fragment approach to calculating changes in atomic charges—the motion of atomic boundaries is included in the calculation. Since ZFSs exist around any gradient bundle, not just atomic basins, studying the movement of the entire gradient field of  $\rho(\mathbf{r})$  provides a more detailed picture of electron rearrangement. The motion of ZFSs between and within atomic basins shows where electron density changes when a chemical reaction

occurs, providing similar results as frontier orbital theory.

## CHAPTER 5

### APPLICATIONS TO ENZYME DESIGN

Rational design of enzymes with novel and enhanced activity is an emerging field combining computational and experimental techniques. Enzyme design began as a purely experimental practice, using methods such as directed evolution and the creation of antibodies. Rational enzyme design involves making specific changes to existing enzymes in an attempt to catalyze reactions involving non-native substrates or to alter the catalytic rate. In this chapter, I discuss applications of QTAIM and bond bundles to two rational design projects of metalloenzymes: (1) the redesign of carboxypeptidase A to catalyze a non-native substrate, and (2) the metal specificity of histone deacetylase.

Metalloenzymes catalyze vital reactions including oxidative metabolism of drugs (cytochrome P450) [122], hydrolysis of DNA and RNA (staphylococcal nuclease) [123], and methane production and oxidation (methyl-coenzyme M reductase) [124]. Unfortunately, metal ions present challenges for computationally modeling enzymes including determining correct metal coordination and spin states, the need for large basis sets in QM calculations, and parameterization of force fields for the metal in molecular dynamics (MD) simulations. Furthermore, noncovalent interactions in the active site are known to play a key role in enzyme activity [125], but our understanding of these interactions and ability to computationally determine their strength is not as advanced as with typical covalent bonds. Hydrogen bonding is often important for properly orienting the substrate for reaction and stabilizing charges during the transition state [126]. Other van der Waals interactions also play a role in enzyme catalysis, and the entire electrostatic environment of the active site has been proposed as being the key reason that enzymes are such efficient catalysts [127, 128]. This phenomenon is known as electrostatic preorganization and will be discussed in more detail in regards to an ongoing project in Chapter 6.

QTAIM provides a method for examining all of these complex bonding features in the active sites of enzymes. However, the use of QTAIM in understanding the mechanisms of enzymes is very recent and the application of topological studies to the enzyme design process is even more rare. Most often, when QTAIM is used for enzyme studies, it is used to predict inhibitors for drug design by evaluating properties at bond critical points [129–131]. To the best of my knowledge, the only direct application of QTAIM to enzyme design was a study of potential hydrogen bonding interactions in a mutant of *Bacillus subtilis* lipase performed by Lin *et al.* [132]. Here I present the application of QTAIM to two metalloenzyme design projects, as well as the first ever application of bond bundle analysis to biological systems. While QTAIM is a useful tool for modeling and designing enzymes, it does not always fully recover the bonding character in enzyme active sites, necessitating the development of more refined chemical bonding models such as gradient bundle analysis.

### 5.1 QM/DMD Methods

Both of the following DFT studies use structures initially obtained from QM/DMD simulations, where DMD is short for discrete molecular dynamics [133–137]. DMD is a simplified method of molecular dynamics using square well potentials to describe interactions. This allows for faster sampling and enables one to efficiently capture dynamics on the order of tens of nanoseconds, in conjunction with the quantum mechanical description of the active site. It has been shown to perform exceptionally well for recapitulating and recovery of metalloenzyme structures, including subtle structural details in active sites [133, 138–140]. QM/DMD is especially useful for metalloenzymes compared to other methods since the metal is treated purely quantum-mechanically, meaning there is no need to parameterize the force field for the metal.

The enzymes are divided into three regions: QM-only, DMD-only, and a shared QM-DMD region. DMD is used to sample the entire protein except for the metal and its immediately coordinated shell, which is held frozen during the sampling. During the DMD phase, the protein undergoes a simulated annealing procedure followed by DMD sampling at a constant

temperature. The DMD parameters for the QM-DMD region are adjusted on-the-fly, based on the QM calculations. Solvent effects are included implicitly in the force field of DMD.

After each DMD sampling step, the lowest energy structure proceeds to the QM step. The QM region is geometrically optimized, allowing all atoms in the QM-only region to move. Atoms at the QM-DMD boundary are capped with hydrogens and the boundary atoms along with the capping hydrogens are frozen. DFT single point energy calculations are then performed on the optimized QM structure. The DMD and QM steps are performed iteratively, until both energy and geometry (based on RMSD calculations) are converged. The nuclear positions from the lowest energy QM step are used for single point active site DFT calculations and the resulting charge density is used for topological analysis.

## 5.2 Predictive Methods for Computational Metalloenzyme Redesign - A Test Case with Carboxypeptidase A

Modified from a manuscript submitted to *Physical Chemistry Chemical Physics*.

Crystal E. Valdez<sup>1</sup>, Amanda Morgenstern<sup>2</sup>, Mark E. Eberhart<sup>3</sup>, and Anastassia N. Alexandrova<sup>4</sup>. 2016.

### 5.2.1 Introduction

Carboxypeptidase A (CPA) is an exopeptidase that preferentially cleaves C-terminal aliphatic and aromatic amino acids from dietary proteins. In this study we predict mutations to CPA that could allow the enzyme to be active toward a non-native substrate, hippuryl-L-aspartate. Figure 5.1 shows the active site of CPA, which consists of a  $\text{Zn}^{2+}$  ion coordinated to two histidines (His69 and His196), a glutamate residue (Glu72), and one water molecule. The zinc ion acts as a Lewis acid, lowering the pKa of the bound water molecule to make it a more potent nucleophile, and polarizing the carbonyl oxygen of the substrate (in this case, hippuryl-L-phenylalanine (HPA)). Key components of an active mutant enzyme include

---

<sup>1</sup>Primary researcher and author

<sup>2</sup>Performed QTAIM analysis

<sup>3</sup>Provided guidance on QTAIM analysis

<sup>4</sup>Corresponding author

proper positioning of the substrate with respect to the catalytic Zn and important residues in the active site, resulting in favorable energetics of the rate determining step.

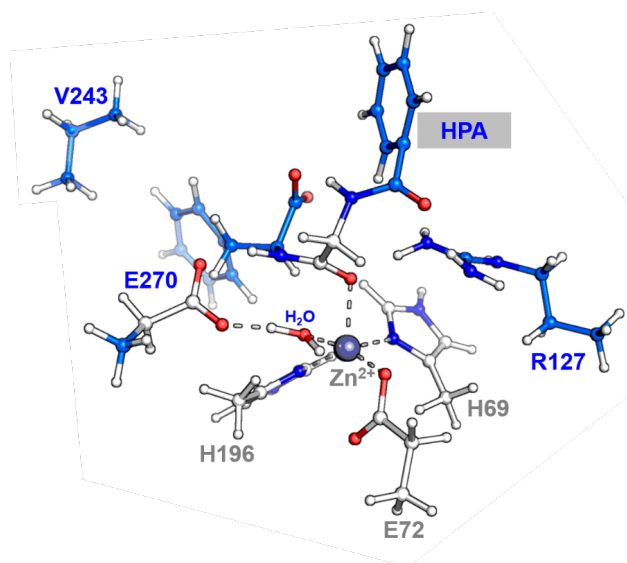


Figure 5.1: Active site of CPA showing partitioning for QM/DMD calculations: grey-QM/DMD region, blue-QM region. Single point DFT calculations were performed on the entire active site pictured here.

While the exact mechanism of CPA is debated, a water-promoted mechanism (Figure 5.2) is believed to be the most energetically favorable [141, 142]. The C-terminal peptide specificity of CPA comes from the highly organized hydrogen bond network within the binding pocket that includes Arg127. Additionally, Arg127 interacts not only with the terminal carboxylic acid, but also with the carbonyl oxygen of the peptide backbone coordinated to zinc, an interaction believed to contribute to catalysis and substrate binding [143]. During the course of the reaction, Arg127 stabilizes the oxyanion developed on the tetrahedral intermediate. Glu270 is also thought to play a vital role in catalysis by providing a hydrogen bond acceptor for the zinc-bound water. Our goal in the design process is to preserve these valuable interactions.

A mutant enzyme (V243R) is computationally predicted to catalyze hydrolysis of the non-native substrate. However, the mutant enzyme is predicted to be less catalytically efficient than the native enzyme based on a higher calculated activation energy barrier (13.3



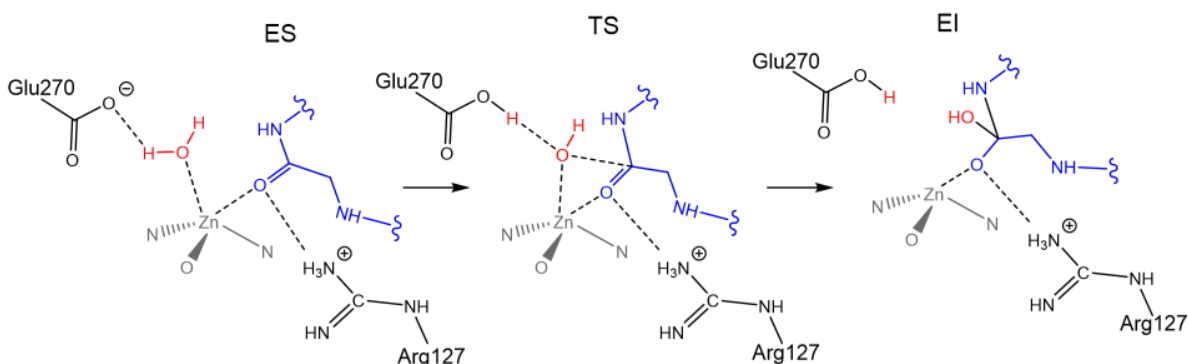


Figure 5.2: The proposed water-promoted mechanism of peptide hydrolysis in CPA (the peptide substrate, shown in blue, is truncated for clarity). The water molecule is shown in red.

$kcal/mol$  for the native enzyme vs.  $21.6 kcal/mol$  for the mutant). Here I present the charge density topology analysis of both native CPA and V243R, rationalizing why the native enzyme should be a more efficient catalyst.

### 5.2.2 Methods

The initial structure of CPA was obtained from the Protein Data Bank (code: 6CPA) [144]. The coordinates from this file were also used to create the V243R mutant enzyme. The crystal structure contains a bound inhibitor, which served as a scaffold to build the substrates for both systems. QM/DMD was performed using the QM and QM-DMD regions shown in Figure 5.2 for dynamic sampling for both the native and mutant enzyme. After annealing, protein sampling occurred at a constant temperature of  $0.10 kcal/(mol * k_B)$  (approximately 50K). QM/DMD convergence was reached after 100 iterations, corresponding to around 50 ns of dynamics. The QM steps were performed with *Turbomole* using a TPSS functional, double  $\zeta$  quality basis set for all atoms except the metal which utilized a def2-TZVPP basis set [145–149]. The Conductor-like Screening Model (COSMO) continuum solvent model was applied with a dielectric constant of 20.0 [150]. A full mechanistic study using the converged active site geometry was performed in *Turbomole*. The rate determining step for the peptide hydrolysis mechanism shown in Figure 5.2 is predicted to have an activation barrier of 13.3

*kcal/mol* for the native enzyme and 21.6 *kcal/mol* in the mutated enzyme.

The electronic charge densities of the stationary points resulting from a QM mechanistic study for the native enzyme and V243R mutant with the bound substrates were calculated with the Amsterdam Density Functional Package (ADF) version 2014 [20, 68, 151] using similar computational parameters as in *Turbomole*. This includes the enzyme-substrate (ES), first transition state (TS1), and enzyme-intermediate (EI) structures. A TPSS functional and COSMO solvent model with a dielectric of 20.0 were utilized [152, 153]. A double  $\zeta$  quality basis set, DZP, was employed for all atoms except the metal, which was calculated using a triple  $\zeta$  quality basis set, TZP [116]. The Bondalyzer add-on package in Tecplot was used to analyze the calculated charge densities [79].

### 5.2.3 Results and Discussion

Two key components to the catalytic activity of CPA are correct positioning of the water molecule for nucleophilic attack and stabilization of the partially charged substrate carbonyl. Figure 5.3 shows bond paths of interest for the ES complexes. In both the native CPA and V243R mutant, the water molecule is positioned near the substrate carbonyl due to coordination with the Zn ion and hydrogen bond donation to E270. The native enzyme has additional hydrogen bonding from the second water hydrogen to E72. In V243R, E72 is doubly coordinated to the Zn ion, rather than acting as a hydrogen bond acceptor for the water molecule. The donation of the second hydrogen atom in the native enzyme results in a more negative Bader charge on the water oxygen (see Table 5.1). This should make the native water molecule a better nucleophile as is indicated by the lower activation energy in the native enzyme as compared to V243R.

The partially negative carbonyl oxygen on the substrate is stabilized through hydrogen bonding to R127 and coordination to the Zn center in both enzymes. The hydrogen bond from R127 to O1 in the reactants is not as strong in V243R as compared to the native enzyme. This is indicated by the charge density at the  $H_{R127-O1}$  bond CP, which has decreased from

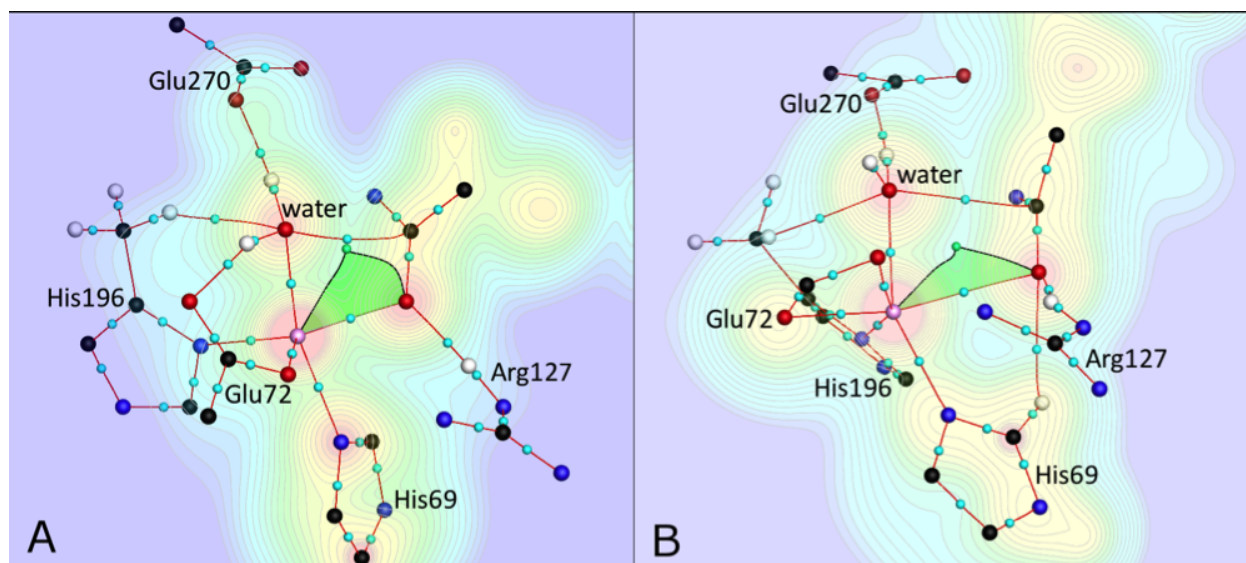


Figure 5.3: Bond paths of interest in the native CPA (A) and V243R mutant (B) ES complexes. Contours in  $\rho(\mathbf{r})$  are drawn on a cut plane on a logarithmic scale from  $10^{-3} - 1 \text{ e/bohr}^3$ . Red lines indicate bond paths. The pictured portion of the Zn–O1 bond bundles are shaded green with black lines showing approximate edges. The following coloring scheme is used: Zn-purple, O-red, C-black, H-white, N-blue, bond CP-cyan, ring CP-green.

0.0487 in the native enzyme to  $0.0128 \text{ e/bohr}^3$  in the mutant (Table 5.2). Additionally, the O1–H<sub>R127</sub>–N<sub>R127</sub> angle has decreased from almost linear in CPA,  $175^\circ$ , to  $147^\circ$  in V243R. However, there is an additional weak hydrogen bonding interaction between the carbonyl oxygen and H69 in the mutant enzyme.

Similarly, the native enzyme has greater stabilization of the carbonyl oxygen from the Zn ion as compared to V243R. This can be seen in the amount of charge density at the Zn–O1 bond CP and also in the size of the Zn–O1 bond bundle. The charge density decreases from 0.0575 in the native enzyme to  $0.0183 \text{ e/bohr}^3$  at the bond CP in V243R. A cut plane of a portion of the bond bundle in the plane of the Zn–O1–ring CP is shown in Figure 5.3. These points define three vertices of the Zn–O1 bond bundle. The pictured portion of the bond bundle in the native enzyme is significantly larger than in V243R, indicating a more stabilizing bonding interaction.

Table 5.1: Bader charges of atoms participating in the tetrahedral intermediate in the CPA hydrolysis mechanism.  $O_{\text{wat}}$  is the water oxygen and C1 and O1 are the carbon and oxygen atoms on the substrate carbonyl.

	ES		TS1		EI	
	Native	Mutant	Native	Mutant	Native	Mutant
$O_{\text{wat}}$	-1.181	-1.135	-1.014	-1.038	-1.003	-1.003
C1	1.120	1.162	1.100	1.080	1.117	1.073
O1	-1.087	-1.103	-1.102	-1.102	-1.089	-1.085
Zn	1.314	1.308	1.297	1.304	1.250	1.262

Table 5.2: Charge density at bond and ring critical points pictured in Figure 5.3 for the rate-determining step of peptide hydrolysis in native CPA and the V243R mutant.

	ES		TS1		EI	
	Native	Mutant	Native	Mutant	Native	Mutant
C1- $O_{\text{wat}}$	-0.0166	0.0155	0.0953	0.0817	0.217	0.204
C1-O1	0.352	0.368	0.338	0.342	0.301	0.298
Zn-O1	0.0575	0.0183	0.0699	0.0599	0.0762	0.0746
Zn- $O_{\text{wat}}$	0.0558	0.0596	0.0327	0.0395	n/a	n/a
$H_{R127}$ -O1	0.0487	0.0128	0.0596	0.0205	0.0412	0.0287
Ring CP	0.0166	0.0129	0.0262	0.0269	n/a	n/a
$H_{H69}$ -O1	n/a	0.0157	n/a	0.0160	n/a	0.0127

In both native CPA and V243R mutant, the Zn ion and hydrogen bond network are able to stabilize the partial charge on the substrate carbonyl. Throughout this step of the reaction, the charge on the oxygen varies by only 0.178 electrons and 0.132 electrons for the native and mutated enzyme, respectively. Overall, the hydrogen bonding networks in both the native and mutated enzyme promote the reaction. The higher activation energy in V243R may be partially due to less stabilizing interactions from both hydrogen bonding and the coordination of the substrate carbonyl to the Zn ion.

In the rate-determining step of the reaction, the  $\text{Zn-O}_{\text{wat}}$  bonding interaction is broken, opening the tetrahedral ring pictured in Figure 5.3. Topologically, the ring opening is achieved by the ring CP in the tetrahedral intermediate moving towards the  $\text{Zn-O}_{\text{wat}}$  bond CP in the TS (see Figure 5.4) until it ultimately annihilates the bond CP in the EI complex. While the ring CP in V243R is geometrically closer to the  $\text{Zn-O}_{\text{wat}}$  bond CP in the ES complex, the charge density region between the ring and bond CP in the native enzyme is “flatter”. The change in  $\rho(\mathbf{r})$  between the ring-bond CPs is only  $0.0392 \text{ e/bohr}^3$  in the native enzyme while it is  $0.0467 \text{ e/bohr}^3$  in V243R. In accordance with the electron-preceding picture [61, 91–93], it should therefore be less energy intensive for the native enzyme ring CP to annihilate the bond CP than in the mutant enzyme, which agrees with our mechanistic predictions.

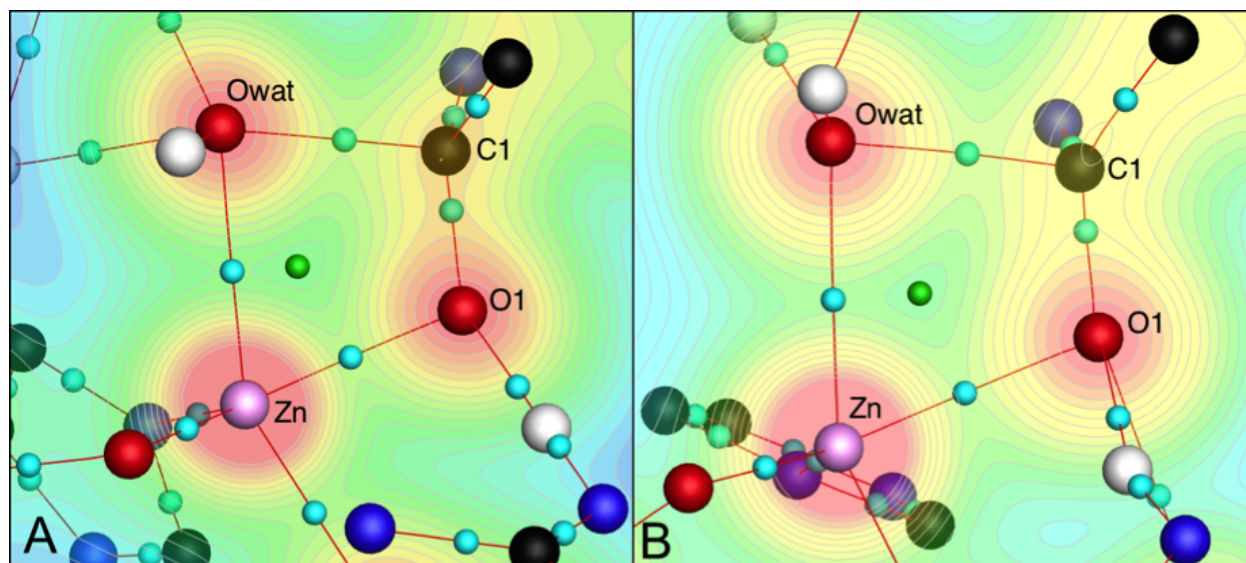


Figure 5.4: Bond paths and critical points for the both the native CPA (A) and V243R mutant (B) TS. The ring CP (green) has moved towards the  $\text{Zn-O}_{\text{wat}}$  bond CP in both enzymes. After the transition state, the ring CP converges with the  $\text{Zn-O}_{\text{wat}}$  bond CP, causing the two critical points to be topologically annihilated and the  $\text{Zn-O}_{\text{wat}}$  bonding interaction to break.

The observed correspondence between values of  $\rho(\mathbf{r})$  at bond and ring CPs in the reacting region combined with bond bundle analysis and the energetic barrier to the reaction outlines a novel strategy for predicting enzyme reactivity. Differences in  $\rho(\mathbf{r})$  in the reactant states of

related enzymes, for example designed computationally to catalyze the same reaction, could serve as a probe of relative catalytic activity prior to the mechanistic study. We believe this probe is very sensitive and non-empirical, responding to the embedding in the protein charges, i.e. the electrostatic pre-organization. Importantly, analysis can be performed on the reactant state alone, prior to performing more costly and often tedious transition state calculations during the mechanistic study. Further in-depth investigation of the utility of this parameter for computational enzymology is discussed in section 6.6.

### 5.3 Metal Specificity of Histone Deacetylase

Modified from a paper published in *Journal of Physical Chemistry B*<sup>1</sup>.

M. R. Nechay<sup>2</sup>, N. M. Gallup<sup>3</sup>, A. Morgenstern<sup>4</sup>, Q. A. Smith<sup>5</sup>, M. E. Eberhart<sup>6</sup>, A. N. Alexandrova<sup>7</sup>. 2016. DOI: 10.1021/acs.jpcb.6b00997.

#### 5.3.1 Introduction

Histone deacetylases (HDACs) are responsible for the removal of acetyl groups from histones, which results in gene silencing (21). Overexpression of HDACs is associated with cancer, and their inhibitors are of particular interest as chemotherapeutics [154–159]. HDAC8 is the most studied HDAC, and has traditionally been considered to be a Zn-dependent enzyme. However, recent experimental assays have challenged this assumption and shown that HDAC8 is catalytically active with a variety of different metals, and that it may be an Fe-dependent enzyme *in vivo* [160]. Here we study the mechanism of HDAC8 utilizing a series of divalent metal ions in physiological abundance ( $\text{Zn}^{2+}$ ,  $\text{Fe}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ni}^{2+}$ , and  $\text{Mg}^{2+}$ ). Two important factors for evaluating the metal dependency of HDAC8 are (1) acti-

---

<sup>1</sup>Reproduced in part with permission from *Michael R. Nechay, Nathan M. Gallup, Amanda Morgenstern, Quentin A. Smith, Mark E. Eberhart, and Anastassia N. Alexandrova. J Phys. Chem. B. Article ASAP, DOI: 10.1021/acs.jpcb.6b00997. Copyright 2016 American Chemical Society.*

<sup>2</sup>Primary researcher and author

<sup>3</sup>Performed calculations and analysis for active site models

<sup>4</sup>Performed QTAIM analysis

<sup>5</sup>Performed preliminary QM/DMD study

<sup>6</sup>Provided guidance on QTAIM analysis

<sup>7</sup>Corresponding author

vation energies of the rate-limiting step with each metal cation, and (2) the binding affinity of the protein for each metal. In this section I present the topological analysis of the charge density in an ongoing effort to explain the calculated and experimentally observed energetics in HDAC8 with each metal ion.

The active site of HDAC8 contains a metal-binding center coordinated to one histidine (H180) and two aspartate (D178 and D267) residues (see Figure 5.5). A water molecule is also present in the crystal structure and interacts with the metal center and two additional histidine residues, H142 and H143. These histidines form dyads with the aspartates D176 and D183, respectively. In the crystal structure of a Y306F mutant (PDB code 2V5W), the acetyl-lysine substrate appears coordinated to the zinc, with a carbonyl oxygen.

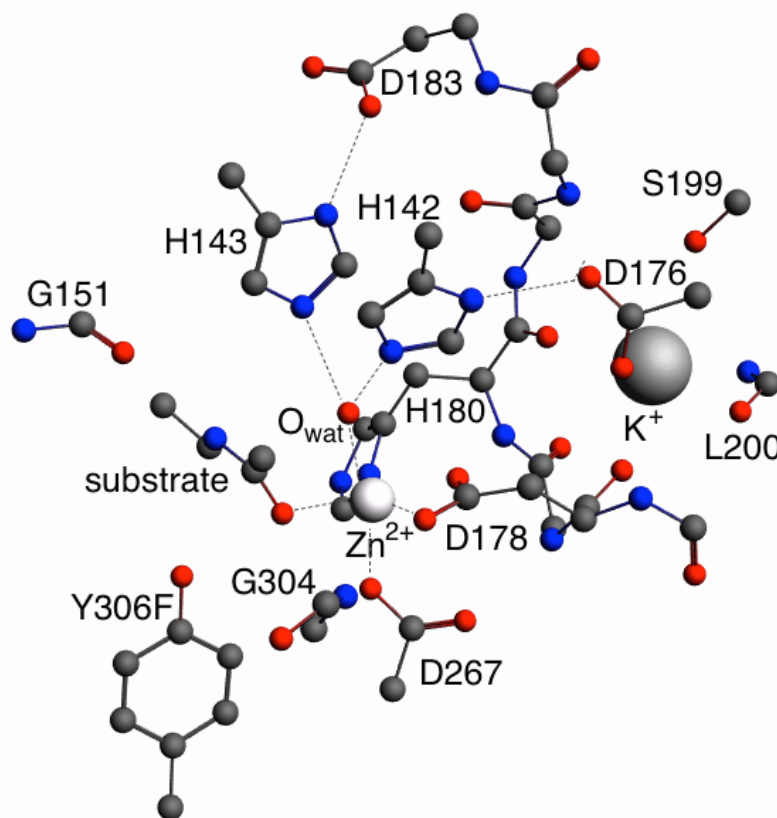


Figure 5.5: The active site of HDAC8 extracted from the crystal structure, highlighting the most critical residues relevant to catalysis within the active site. In this work, F306 was mutated to tyrosine to facilitate catalysis. This full structure was used as the QM region in QM/DMD calculations and for the QM mechanistic study and charge density calculations.

We found a proposed proton shuttle mechanism [161] to be the most viable through detailed QM/DMD calculations [162]. In the proton shuttle mechanism (Figure 5.6), H142 and H143 are initially singly protonated. H143 alternates between acting as a base and acting as an acid, first abstracting a proton from water and subsequently transferring it to the substrate nitrogen. This results in two transition steps during the process of cleaving the C-N bond in the substrate, which we refer to as ts1 and ts2. The reaction coordinate for the proton shuttle mechanism with all six metal ions was calculated using the methods described in Section 5.3.2.

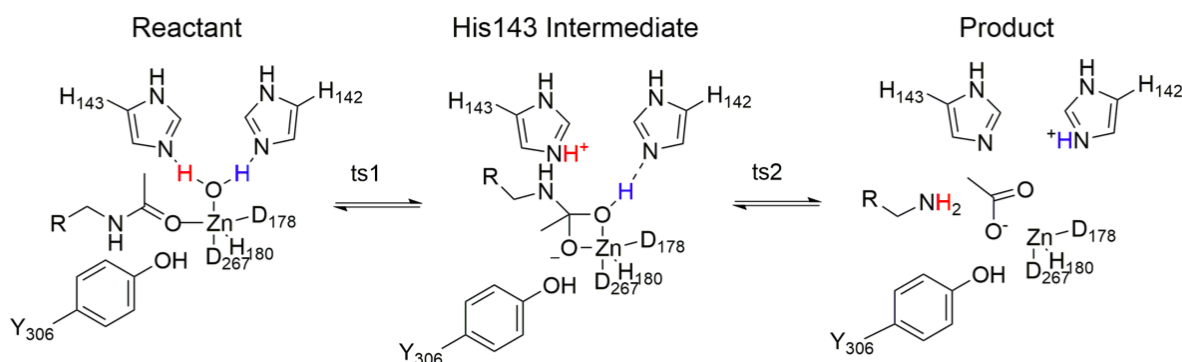


Figure 5.6: Proposed proton shuttle mechanism for HDAC8.

For the second part of this study, we explored how metal binding affinity may affect the overall reactivity of HDAC8. Gantt and coworkers have shown that several of the metal ions in this investigation were inactive *in vitro* [160]. However, many of our computed rate-limiting transition state energies are well under the ceiling of catalytic inactivity. It is possible that the binding affinity of each metal towards HDAC8 could have an effect on *in vitro* catalytic activity, and may play an important role for *in vivo* metal selection. Here, we attempt to elucidate the significance of this effect using metal swapping QM calculations.

The combination of  $\Delta\Delta G$  data for metal swapping with the catalytic rate constants from mechanistic studies is required to reproduce the experimentally observed trend in metal-dependent performance. We predict  $\text{Co}^{2+}$  and  $\text{Zn}^{2+}$  to be the most active metals in HDAC8, followed by  $\text{Fe}^{2+}$ , and  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  to be the least active.  $\text{Ni}^{2+}$  proves to be an interesting



outlier for which we recommend further study. We show that results from the metal swapping portion of this study can be quantitatively rationalized using QTAIM analysis on the charge densities for each metal-substituted active site. For the activation energy study, however, QTAIM is only able to provide a qualitative explanation for energetic results, presenting a clear case where higher resolution bonding models, such as gradient bundle analysis, may be required.

### 5.3.2 Methods

The initial structure of HDAC8 was obtained from the Protein Data Bank crystal structure of the Y306F mutant (code: 2V5W) [163]. Residue 306 was transformed from phenylalanine to tyrosine using UCSF Chimera. The protonation states of all residues were chosen in accord with their pKa values at neutral pH, including singly protonated His180 on the  $\delta$  site, and singly protonated H142 and H143 on the  $\epsilon$  site.

An appropriate choice of QM and DMD subsystem sizes was critical for QM/DMD simulations. A QM region of 165 atoms was found to be necessary in order to produce logical results (such as binding of the substrate) and is pictured in Figure 5.5. QM calculations were performed using DFT as implemented in *Turbomole* v6.5 [145]. Solvent effects were included implicitly in the force field of DMD. In the QM region, solvent was modeled using COSMO [150] with a dielectric constant of 4.0 (in consideration of the buried active site). Separate calculations were also performed using non-polarizable point charges (AMBER force field), but no significant changes in the mechanism were observed. (The results of these calculations are reported in the supplementary information of [162]). All reported charge densities and energies are therefore from the COSMO calculations. DFT single point energy calculations were carried out using the TPSS functional [146, 147]. The def2-TZVPP basis set [149] was used for the metal atoms, while all other atoms were described by def2-SVP [149]. The Grimme dispersion correction [164] was included for all QM calculations.

The lowest energy structures from QM/DMD proceeded to the QM mechanistic study. All calculations of the stationary points on the reaction profile were optimized with *Turbomole*

using the TPSS functional and Def2-TZVPP basis set for all metals, and Def2-SVP basis sets for non-metal atoms. Vibrational frequency analysis was done using the same level of theory, and the nature of all transition states was determined by the presence of a single imaginary frequency aligned with the reaction coordinate. Single point energy calculations were carried out on all stationary points using the larger Def2-TZVPP basis set for all atoms. The full reaction coordinate for all metals, including zero point energy and thermal and entropic corrections, is shown in Figure 5.7.

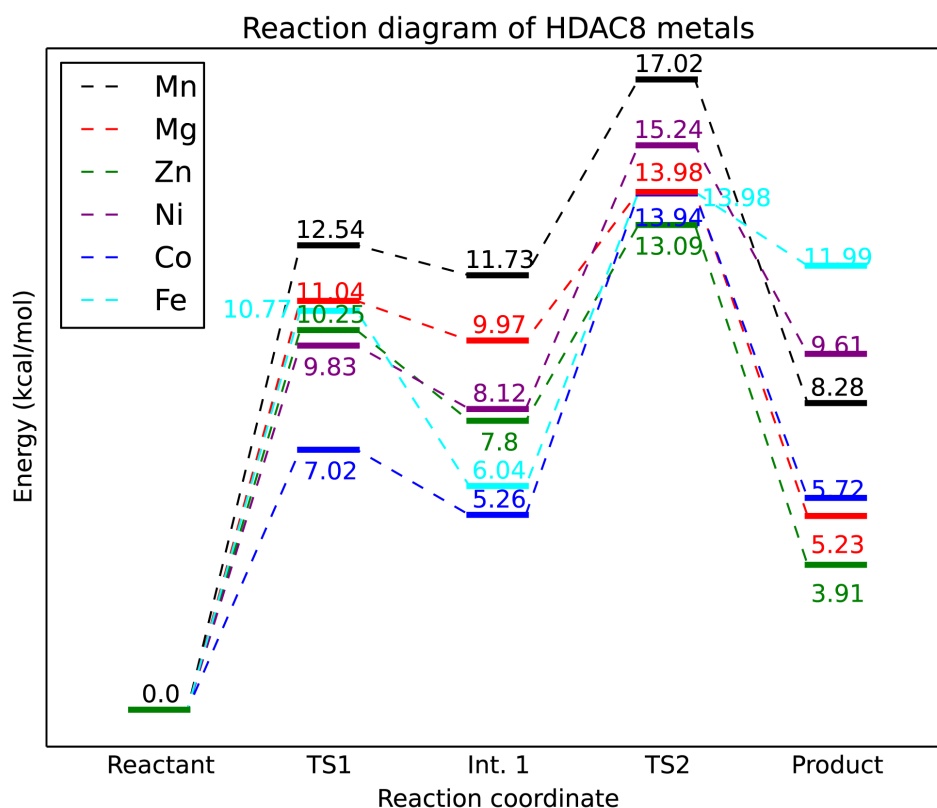


Figure 5.7: Calculated reaction pathway for HDAC8 with a variety of divalent metal ions.

Born-Oppenheimer molecular dynamics (BOMD) simulations were done for the full DFT sampling of the complexes of the studied metals with organic chelators in order to assess the relative metal binding affinities to the HDAC8 protein. To calculate the  $\Delta\Delta G$  for metal swapping, we first compared the accuracy of DFT to experiment using the experimentally

known stability constants of the complexes of the studied metals with GEDTA and DTPA [165]. From these stability constants, the free energies of reactions (1-4) in Figure 5.8 are known. Further, from  $\Delta\Delta G$  of reactions (1) and (2), one can close the thermodynamic cycle for DTPA on the upper left in Figure 5.8. Analogously, one can close the cycle for GEDTA on the upper right in Figure 5.8, using  $\Delta\Delta G$  of reactions (3) and (4). Through combination of these two cycles, one can then calculate the  $\Delta G$  of metal swapping between the two chelators,  $\text{DTPA-M}_a^{2-} + \text{GEDTA-M}_b^{2-} \rightarrow \text{DTPA-M}_b^{2-} + \text{GEDTA-M}_a^{2-}$ , bypassing the complicated calculation of the solvated metal ions,  $\text{M}_a^{2+}(\text{aq})$  and  $\text{M}_b^{2+}(\text{aq})$ .

The  $\Delta G$  values of metal swapping were calculated both from the experimental data, and with DFT (See Appendix B for results). The structures of the metal-GEDTA and metal-DTPA complexes were geometry-optimized and also subjected to 5 ps of BOMD, to verify the ground state geometries. To avoid additional desolvation complications, polydentate ligation of the chelating agent was selected such that the bound metal had no solvent access, preventing explicit water coordination, and allowing for the exclusive use of COSMO. The computed  $\Delta G$  values were found to be in good agreement with experiment, having an average error of  $0.5 \pm 1.5 \text{ kcal/mol}$ .

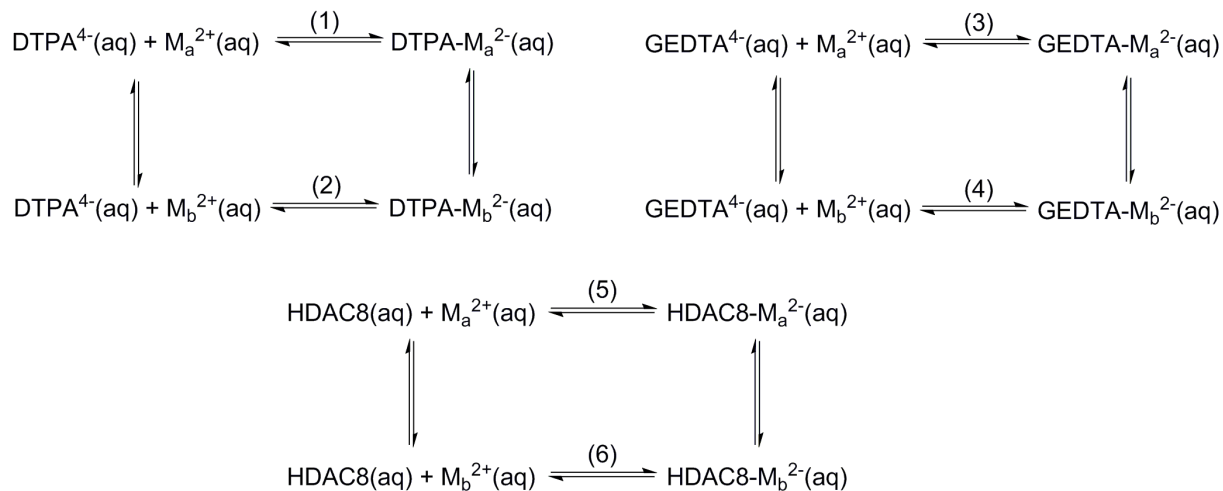


Figure 5.8: Schematic of all considered thermodynamic cycles exploited for relative metal binding affinities in HDAC8.

This methodology was then applied to a theoretical metal swap between DTPA and HDAC8:  $\text{DTPA-M}_b^{2-} + \text{HDAC8-Co}^{2-} \rightarrow \text{DTPA-Co}^{2-} + \text{HDAC8-M}_b^{2-}$ , using  $\Delta G$  of reactions (1) and (2), and (5) and (6) in Figure 5.8, and with  $M_a$  set to  $\text{Co}^{2+}$  as a reference. The above reaction can be paired with the experimental  $\Delta G$  of reactions to yield  $\Delta G$  of the following process:  $\text{M}_b^{2+}(\text{aq}) + \text{HDAC8-Co}^{2-} \rightarrow \text{Co}^{2+}(\text{aq}) + \text{HDAC8-M}_b^{2-}$ . This gives the desired relative affinities of the different metals to HDAC8, collected in Table 5.3.

Table 5.3:  $\Delta\Delta G$  of binding between metal ions and HDAC8, relative to  $\text{Co}^{2+}$ . Energies are in *kcal/mol*.

Metal	$\Delta\Delta G$ binding
$\text{Co}^{2+}$	0.00
$\text{Zn}^{2+}$	1.99
$\text{Fe}^{2+}$	3.52
$\text{Mn}^{2+}$	6.28
$\text{Mg}^{2+}$	16.75
$\text{Ni}^{2+}$	-4.33

The charge densities for the reactant geometries of HDAC8 with all 6 metal ions were calculated with the Amsterdam Density Functional Package (ADF) version 2014.01 [20, 68, 151] using similar computational parameters as in the QM mechanistic study. A TPSS functional [147, 166] and COSMO solvent model with a dielectric constant of 4.0 were utilized. A double  $\zeta$  quality basis set, DZP, was employed for all atoms except the metal, which was calculated using a triple  $\zeta$  quality basis set, TZP [116]. Spin-unrestricted calculations were performed on the Zn and Mg active sites, while all other metals were calculated in their high-spin states. The Bondalyzer add-on package in Tecplot [79] was used to analyze the resultant charge densities.

### 5.3.3 Results and Discussion

The reaction coordinate in Figure 5.7 shows the second transition state (ts2) to be rate limiting in all cases. This result marks a pronounced deviation from previously published models, which claimed the first transition state, ts1, to be rate-limiting [161, 167] for  $\text{Zn}^{2+}$ . This may be an artifact of not using a large enough QM region in the earlier studies. We further find that  $\text{Zn}^{2+}$  facilitates the fastest catalytic path, followed by  $\text{Co}^{2+}$ ,  $\text{Fe}^{2+}$  and  $\text{Mg}^{2+}$ , then  $\text{Ni}^{2+}$ , and finally  $\text{Mn}^{2+}$ . This is in contrast to experimental assays published by Gantt and coworkers [160] who found  $\text{Co}^{2+}$  and  $\text{Fe}^{2+}$  to be the most active, followed by  $\text{Zn}^{2+}$ , and little reactivity for  $\text{Ni}^{2+}$  and  $\text{Mn}^{2+}$ .

The main role of the metals in the reaction comes in the first step, where the carbonyl of the substrate is activated for nucleophilic attack. This determines both the activation barrier and the energy of the resultant intermediate, thereby shaping the entire reaction profile. We examine the reactant states using QTAIM, to shed light on the electronic effects leading to the metal-dependent performance. The lower energy barriers of ts1 for  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Fe}^{2+}$ , and  $\text{Co}^{2+}$  may be in part due to the higher number of occupied d orbitals than for  $\text{Mn}^{2+}$  ( $\text{Mg}^{2+}$  obviously lacking any).

The greater number of d-electrons implies that orbitals higher up in the enzyme’s molecular orbital manifold will be filled. In general, orbitals higher in the manifold are characterized by a greater number of antibonding orbitals and hence more interatomic nodal planes. It is orbitals of this type that stabilize topological rings and cages by contributing density along bond paths though not along nodal planes. This combination leads to deeper rings and greater curvature at bond CPs. In the plane of a ring of nuclei as shown in Figure 5.9, the ring point is a minimum in the charge density. In order for a ring point to exist in this region, a bond path must form between the water molecule and substrate carbonyl.

A bond path is present in the reactant state between the substrate carbonyl and water molecule when  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Fe}^{2+}$ , or  $\text{Co}^{2+}$  is placed in the active site of HDAC8. This bond path is not present for  $\text{Mg}^{2+}$  or  $\text{Mn}^{2+}$ . While this bond is not indicative of a strong

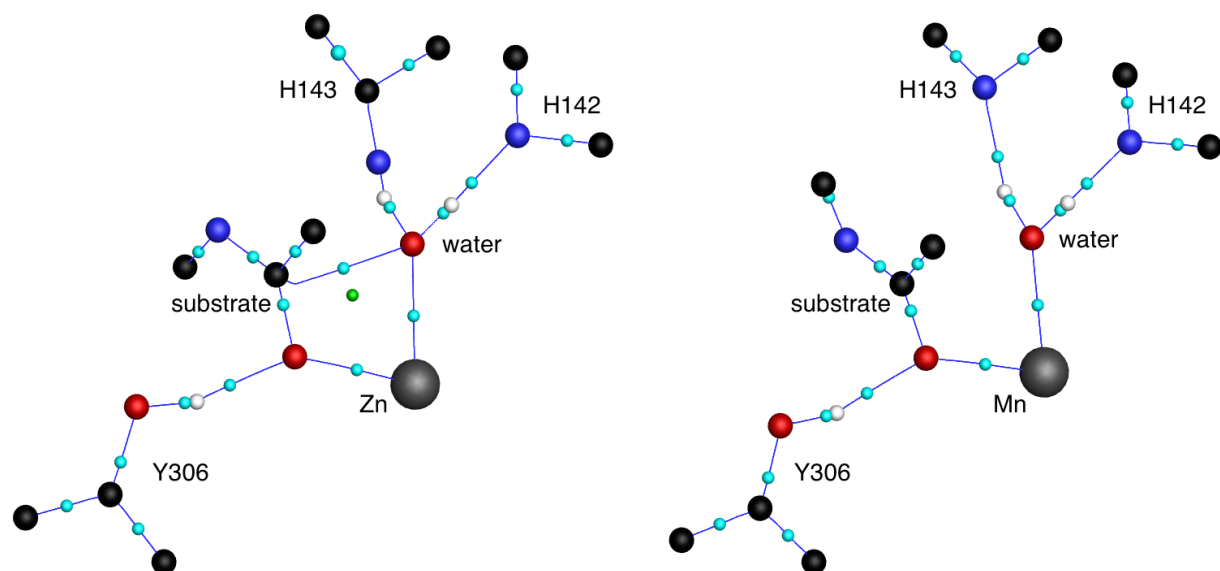


Figure 5.9: Critical points and bond paths of interest in the active site of HDAC8 with Zn (left) and Mn (right). When Zn is present in the active site a bond path forms between the water oxygen and the carbon atom from the substrate carbonyl. This topologically necessitates a ring critical point to exist in the active site as well. When Mn is present, no ring CP is found as there is not a bond path between the water and substrate carbonyl. Ni, Fe, and Co give the same topology as Zn, while Mg has the same topology as Mn. Sphere coloring is as follows: C-black, N-blue, O-red, Metal-grey, bond CP-cyan, ring CP-green.

interaction (based on low values of charge density at the bond CP as well as the curved nature of the bond path), it still indicates a stabilizing interaction between the water molecule and substrate. It can be seen in the proton shuttle mechanism (Figure 5.6) that when a proton is transferred from the water molecule to H143, a bond forms between the carbonyl carbon and water oxygen. This bond has already begun to form in the reactant state for  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Fe}^{2+}$ , and  $\text{Co}^{2+}$ , which facilitates the removal of a water hydrogen and ultimately lowers the activation energy of this reaction step.

However, the correlation of this topological effect with the barrier height only provides a qualitative rationale for activation energies. The ts1 activation energies for  $\text{Zn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Fe}^{2+}$ , and  $\text{Mg}^{2+}$  are all within 1.2 *kcal/mol*, generally considered the margin of error for DFT energy calculations. The presence/absence of the bond path does not provide information to predict that  $\text{Co}^{2+}$  would have a lower ts1 activation barrier by almost 3 *kcal/mol* compared to any other metal tested, or that  $\text{Mn}^{2+}$  would have a higher ts1 barrier of 1.5 *kcal/mol*. Values of  $\rho(\mathbf{r})$  and  $\nabla^2\rho(\mathbf{r})$  at bond and ring CPs also did not correlate to the activation energies. While the existence of a bond path sheds some light on the abilities of various metals to catalyze deacetylation, the lack of quantitateness of this analysis highlights a need for advancements in analyzing the topology of the charge density, which will be discussed in more detail in sections 6.3 and 6.6.

The metal binding affinities in HDAC8, however, can be more fully explained using QTAIM. According to the  $\Delta\Delta G$  calculations described in section 5.3.2,  $\text{Co}^{2+}$  is predicted to bind to HDAC8 very strongly. It is followed by  $\text{Zn}^{2+}$  and  $\text{Fe}^{2+}$ .  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  have considerably smaller affinities for HDAC8.  $\text{Ni}^{2+}$  is calculated to have the highest binding affinity towards HDAC8 among all the studied metals, in contrast to experimental results. This ordering can be related to the active site geometries and corresponding electronic effects. A particularly direct interaction exists between the metal and H180, which is apparently highly important for metal binding. The relationship between binding affinity and the amount of charge density at the  $\text{H}_{\text{H180}}$ -metal bond CP is linear (Figure 5.10) with an  $R^2$  value of 0.976.

The charge density at the bond CP is a consequence of  $\sigma$ -bonding and  $\pi$ -back-donation with the available d-atomic orbitals (AOs) on the given metal. Mg, lacking any occupied d-AOs, has a significantly lower charge density at the bond CP and lower binding affinity than any of the transition metals, despite having a relatively short bond length (which is simply due to its smaller size). Thus, the geometric and electronic parameters of this  $H_{H180}$ -metal bond can serve as predictors of the metal binding affinity in this case.

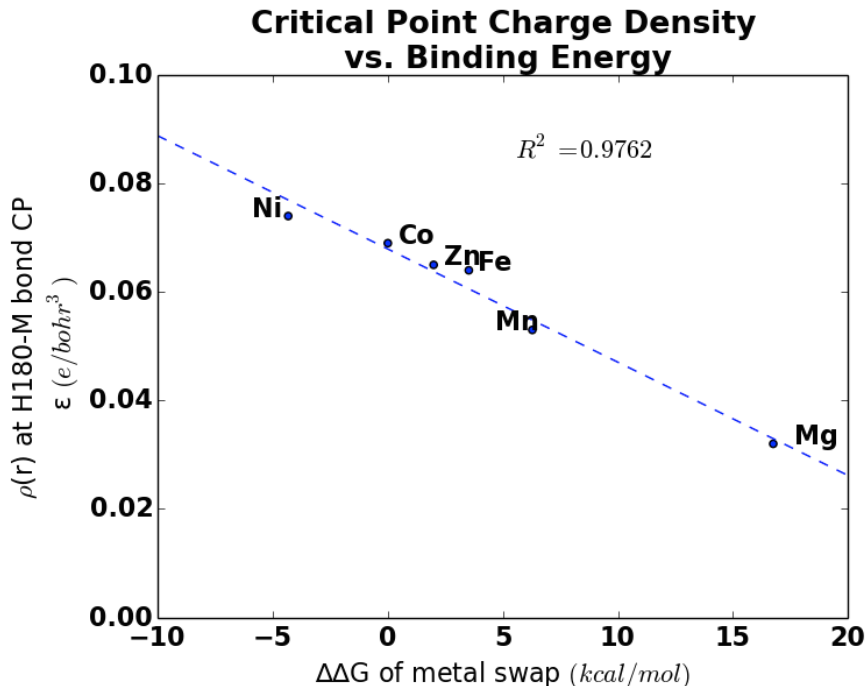


Figure 5.10: Relationship between  $\rho(r)$  at  $H_{H180}$ -metal bond CPs and  $\Delta\Delta G$  of metal swapping.

The difference in metal binding affinity may play a key role in the overall catalytic activity of enzymes such as HDAC8 capable of utilizing different metals, and elucidate its preference for any one particular ion. We evaluate the ability for the protein to retain a metal and subsequently use it for catalysis. Our simple probabilistic model captures the combined effects of metal-binding equilibration and subsequent catalysis relative to any particular ion. We use the equation,

$$k_{rel} = k_{cat} * k_{binding} \quad (5.1)$$



where  $k_{rel}$  stands for the relative total catalytic activity of a particular protein-metal complex,  $k_{cat}$  represents the catalytic activity as determined by QM mechanistic analysis, and  $k_{binding}$  embodies the relative equilibrium binding constant between a metal and an enzyme. These values can be represented as Boltzmann distributions utilizing the above computed quantities in catalysis and binding affinity to yield

$$k_{rel} = \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) * \exp\left(-\frac{\Delta\Delta G_{binding}}{RT}\right) \quad (5.2)$$

whose values are summarized in Table 5.4.

Table 5.4: Predicted and experimental [160] total catalytic activities for different metal ions in HDAC8. Binding affinity energies were taken relative to  $\text{Co}^{2+}$ . Normalized values are relative to the predicted most active metal ion-protein complex.

Metal	$k_{rel}$	$k_{rel}$	$K_{cat} \text{ (s}^{-1}\text{)}$
	Predicted	Normalized	Experimental
$\text{Ni}^{2+}$	1.89E-08	1.00	n/a
$\text{Co}^{2+}$	7.64E-11	4.03E-03	1.2
$\text{Zn}^{2+}$	1.27E-11	6.70E-04	0.90
$\text{Fe}^{2+}$	3.11E-14	1.64E-06	0.48
$\text{Mn}^{2+}$	1.37E-17	7.11E-10	n/a
$\text{Mg}^{2+}$	1.46E-23	7.71E-16	n/a

This model qualitatively predicts the trend seen in catalytic activity for these metals in experiment [160], though the trend is slightly perturbed, either due to experimental errors, or inaccuracies of our model.  $\text{Co}^{2+}$  is predicted as the most active, followed by  $\text{Zn}^{2+}$  and  $\text{Fe}^{2+}$ . Additionally, the experimentally non-catalytic metal  $\text{Mn}^{2+}$  is predicted as being four orders of magnitude less reactive than  $\text{Fe}^{2+}$ . This is due to both manganese's poor binding affinity to the active site of HDAC8 and higher  $\Delta G^\ddagger$ .  $\text{Zn}^{2+}$  is predicted to be less catalytically active than  $\text{Co}^{2+}$ , despite its low reaction barrier, due to its poor affinity, while  $\text{Co}^{2+}$  exhibits high catalytic activity in spite of a higher reaction barrier.  $\text{Fe}^{2+}$  competes with  $\text{Co}^{2+}$  and  $\text{Zn}^{2+}$ , even though it has a higher reaction barrier than both of these cations.

$\text{Ni}^{2+}$ , which was experimentally shown to be inactive, is a clear outlier, and predicted to be exceptionally catalytic, due to high binding affinity. This may be an artifact of the simple model, and requires further investigation. Another possibility is that the coordination of  $\text{Ni}^{2+}$  in HDAC8 may be different from other metals tested here.  $\text{Mg}^{2+}$  has not been tested experimentally, but we found that it has reasonable reaction energetics. However, its affinity to HDAC8 is very low, and  $\text{Mg}^{2+}$  is ultimately predicted by our calculations to have a very low catalytic activity. Thus, it seems that the d-AO structure is required in HDAC8, despite the simple Lewis acid catalysis performed by the metal in this enzyme.

## 5.4 Conclusions

Using QTAIM and bond bundle analysis, we were able to justify the differences in activation energies for native CPA and a V243R mutant for a non-native substrate. The hydrogen bonding network in the native enzyme was seen to be more favorable than the mutant enzyme due to a greater number of hydrogen bond paths with a higher value of  $\rho(\mathbf{r})$  at bond CPs. The Zn ion was also shown to interact and activate the substrate carbonyl in the native CPA more than in V243R. This analysis was based on values of charge density as well as by the size of the Zn–O<sub>1</sub> bond bundle. Finally, the charge density between the ring CP in the tetrahedral intermediate and the Zn–O<sub>wat</sub> bond CP was flatter in the native enzyme, leading to a more energetically facile annihilation of the bond and ring CPs, which is topologically required for the reaction to occur.

In our second enzyme study, we were able to understand the metal binding affinities to HDAC based on the amount of  $\rho(\mathbf{r})$  at bond critical points. However, we were only able to qualitatively rationalize the predicted activation energies of the reaction based on the existence of a bond path (and therefore ring CP) between the water oxygen and substrate carbonyl. This study highlights a need for a higher resolution picture of the charge density to better understand how and why chemical reactions occur. We are currently working to extend gradient bundle analysis techniques to be applicable to non symmetric systems such as enzymes. Details of these improvements to GBA are discussed in the next chapter.

## CHAPTER 6

### FUTURE WORK AND CONCLUSIONS

Thus far, we have shown that gradient bundles can be used to understand valence electron structure, recover bond dissociation energies, and visualize chemical reactivity in small molecules. The next task in developing GBA is to search for structure–property relationships that provide chemical insight into bonding interactions that can not be explained using existing chemical bonding models. One instance where this advancement is needed was highlighted in the previous chapter on metalloenzyme design. While the existence of a bond path correlated to the initial activation energies for deacetylation catalyzed by HDAC with most metals, the iron cation did not follow the predicted trend. In order to further analyze the bonding interactions between the substrate carbonyl and water molecule in HDAC, GBA methods must be adapted to be applicable to more complex systems. We hypothesize that the full gradient field in the active site of HDAC does hold the information necessary to predict which metal cations would make the most efficient catalysis, we just need the methods to extract this information from the charge density.

To this end, the following work is being done to advance gradient bundle analysis. We are quantifying the ZFS picture of reactivity by calculating gradient bundle condensed Fukui functions,  $f_{GB}$ . In addition to viewing the movement of ZFSs, this will enable a change in electron count to be calculated within each gradient bundle. Observing the motion of ZFSs and  $f_{GB}$  highlights a method for predicting which gradient bundles will be most energetically favorable to change their electron count, based only on the differential geometry of gradient paths and ZFSs. Our goal with this project is to answer the almost century old question originally posed by Slater, “where are the HOMO and LUMO in the ground state charge density?”.

Work is also being done to expand GBA to nonlinear systems. We are currently developing a volume integration algorithm to allow 2D plots of electron count and energy to be created around bond paths in Tecplot. This project requires finding meaningful ways to interpret gradient bundle analysis results based on using stereographic projections of data. Once these extensions to GBA are in place we will be able to complete projects studying the gradient bundles in aromatic and anti-aromatic compounds, as well as in solid state systems, such as the ongoing work attempting to elucidate the unique properties of Ir. Finally, we are using QTAIM and GBA to determine the size of active site cluster models required to properly model enzyme mechanisms and kinetics, as well as determining if the charge density in the active site alone can be used to capture electrostatic preorganization in enzymes.

### 6.1 Gradient Bundle Condensed Fukui Functions

The ability to qualitatively analyze reactivity using the movement of ZFSs bounding gradient bundles was demonstrated in chapter 4. This analysis can be made quantitative by additionally calculating the change in electron density in gradient bundles as the electron count in a system changes. The result of this calculation is a gradient bundle condensed Fukui function,  $f_{GB}$ , which provides a more detailed picture of site reactivity than atom condensed Fukui functions. GB condensed Fukui functions also provide evidence of, and explanation for, negative Fukui functions, which is a topic of debate [168–173]. To demonstrate this novel method, here I present  $f_{GB}$  calculations for  $F_2$  and  $N_2$ .

The computational methods for all of the following data was the same as in section 4.5, except that a triple  $\zeta$  basis set was used and no relativistic corrections were made. The RMF method for the condensed Fukui function was integrated over gradient bundles using the equation,

$$f_{GB}^- = \int_{\omega_i} \rho(\mathbf{r}, N) dV - \int_{\omega_i^-} \rho(\mathbf{r}, N - 1) dV \quad (6.1)$$

for the removal of an electron and with

$$f_{GB}^+ = \int_{\omega_i^-} \rho(\mathbf{r}, N+1) dV - \int_{\omega_i} \rho(\mathbf{r}, N) dV \quad (6.2)$$

for the addition of an electron.  $\omega_i$  indicates the integration is performed over any gradient bundle, not just an atomic basin. Note that both equations are set up such that the Fukui function will be positive in the region most likely to undergo nucleophilic attack ( $f_{GB}^+$ ) or electrophilic attack ( $f_{GB}^-$ ).

The gradient bundle condensed Fukui functions for  $F_2$  are shown in the left of Figure 6.1 and Figure 6.2. The plots on the right side of the figures are the percent change in the number of valence electrons in each gradient bundle using

$$f_{GB,\%}^- = 100 \left( \frac{\int_{\omega_i} \rho_V(\mathbf{r}, N) dV - \int_{\omega_i^-} \rho_V(\mathbf{r}, N-1) dV}{\int_{\omega_i} \rho_V(\mathbf{r}, N) dV} \right), \text{ and} \quad (6.3)$$

$$f_{GB,\%}^+ = 100 \left( \frac{\int_{\omega_i^-} \rho_V(\mathbf{r}, N+1) dV - \int_{\omega_i} \rho_V(\mathbf{r}, N) dV}{\int_{\omega_i^-} \rho_V(\mathbf{r}, N+1) dV} \right). \quad (6.4)$$

where  $\rho_V$  indicates valence electron density which is calculated using the methods described in chapter 3. The percent change in electron count may be a more meaningful number than  $f_{GB}$  when using rotational gradient bundles, as it accounts for the fact that gradient bundles near the internuclear axis have a much smaller volume and will therefore always have a smaller change in electron count than the gradient bundles near the center of each atomic basin.

Maxima in  $f_{GB,\%}^-$  for molecular fluorine match the frontier orbital picture of reactivity. Electrophilic attack on  $F_2$  is predicted to occur  $90^\circ$  from the bond path, near the center of each atomic basin, while the Fukui function is negative near the bond path and interatomic surface ( $0^\circ$ ) and pointing away from the bond path ( $180^\circ$ ). This indicates that if an electron is removed from  $F_2$ , charge density will actually be added to the regions with  $-f_{GB}$ . The HOMO of  $F_2$  is shown in the left side of Figure 6.3 and has a lobe in the center of each fluorine atomic basin. There is a nodal plane through the center of the molecule that contains the

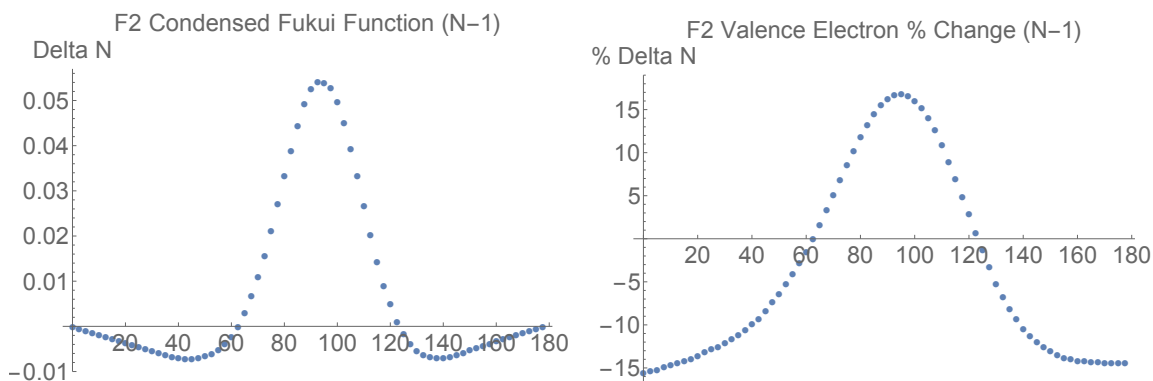


Figure 6.1: Plots of  $f_{GB}^-$  (left) and  $f_{GB,\%}^-$  (right) around one fluorine atom in  $F_2$ . Gradient bundle angles along the x-axes are measured with respect to the bond path, such that the GB at  $0^\circ$  contains the bond path and is bounded by the interatomic surface.

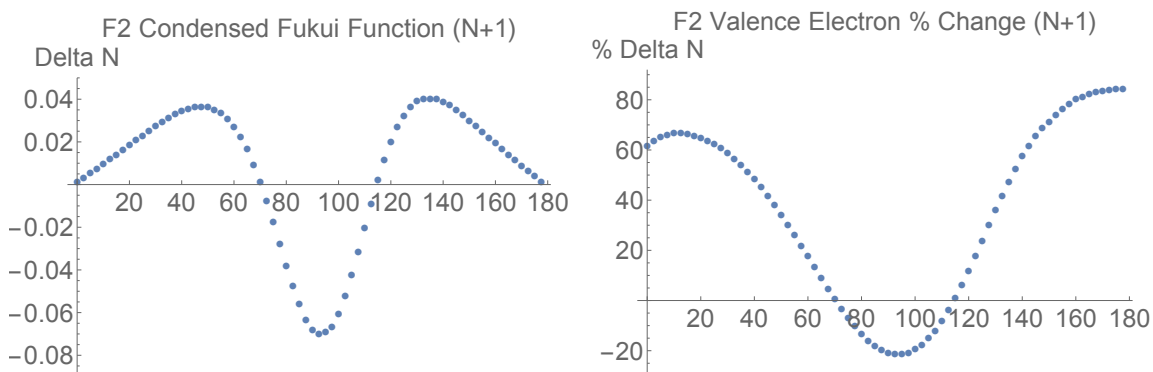


Figure 6.2: Plots of  $f_{GB}^+$  (left) and  $f_{GB,\%}^+$  (right) for one fluorine atom in  $F_2$ .

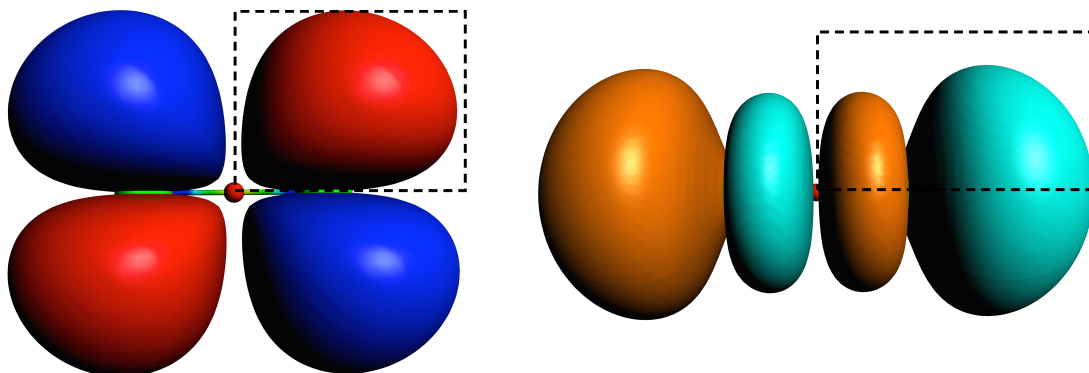


Figure 6.3: The HOMO (left) and LUMO (right) for a neutral fluorine molecule. Black dashed boxes indicate the regions of the orbital plots that geometrically match the Fukui function plots in Figure 6.1 and Figure 6.2.

IAS and bond path. Orbital lobes are present where maxima in  $f_{GB,\%}^-$  occur and nodal planes match the locations of  $-f_{GB,\%}^-$ .

The LUMO of fluorine is pictured on the right side of Figure 6.3 and has two lobes around each fluorine atom, and a nodal plane through the center of the atomic basins, normal to the bond path. This is also recovered in the plot of  $f_{GB,\%}^+$  in Figure 6.2, which shows that electron density will be added to both sides of the atomic basin, and there will be a depletion of charge in the center of the basin,  $90^\circ$  from the bond path. The majority of the charge density will add to the gradient bundles away from the bond path, closer to  $180^\circ$ , which matches where the largest part of the LUMO is located. The nodal plane through the interatomic surface is also recovered as  $f_{GB,\%}^+$  has local minimum at  $0^\circ$ .

Figure 6.4 shows the percent change in valence electron count in gradient bundles for molecular nitrogen. Again, the most reactive sites for electrophilic and nucleophilic attack on nitrogen as predicted by maxima in  $f_{GB,\%}$  match the HOMO and LUMO, respectively (Figure 6.5).  $f_{GB,\%}^-$  has two maxima, one along the interatomic surface ( $0^\circ$ ), and another at  $180^\circ$ , matching the locations of the two lobes of the HOMO. The Fukui function is negative through the center of the atom, matching the nodal plane through the nuclei. There is a single maximum in  $f_{GB,\%}^+$ ,  $120^\circ$  from the bond path, which is in line with the single LUMO lobe in each atomic basin. There are two regions with  $-f_{GB}$ , which match the two orthogonal nodal planes through the bond path and interatomic surface.

Condensed Fukui functions may be particularly useful for locating regions where the Fukui function is negative. It was originally argued that that electron density could not be removed from any regions in a molecule when electrons were added to the system, and vice versa, and therefore negative Fukui functions were not meaningful, but were the result of inaccurate calculation methods [171–173]. In the frozen orbital approximation, this is indeed impossible, as the HOMO and LUMO only show regions where electrons will be lost or gained, respectively. There is no “negative” HOMO or LUMO.

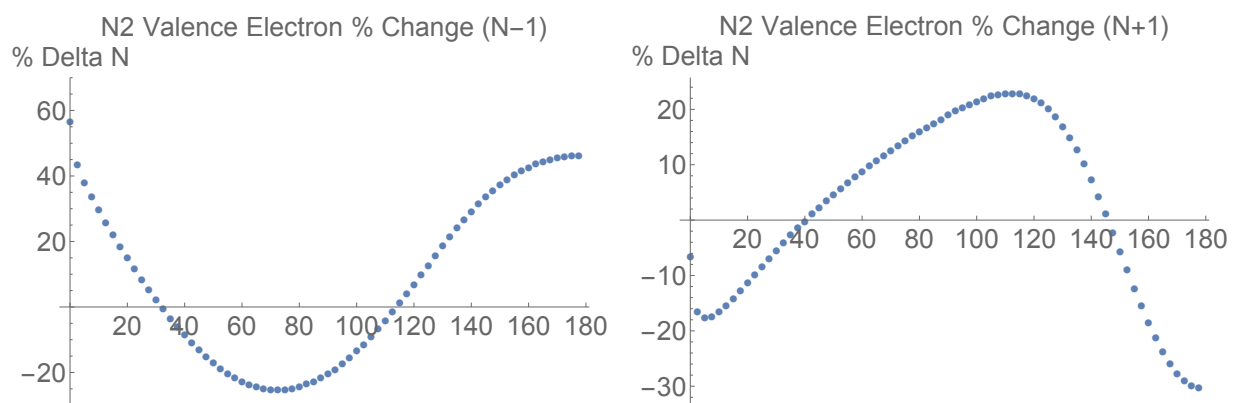


Figure 6.4: Plots of  $f_{GB,\%}^-$  (left) and  $f_{GB,\%}^+$  (right) around one nitrogen atom in  $N_2$ .

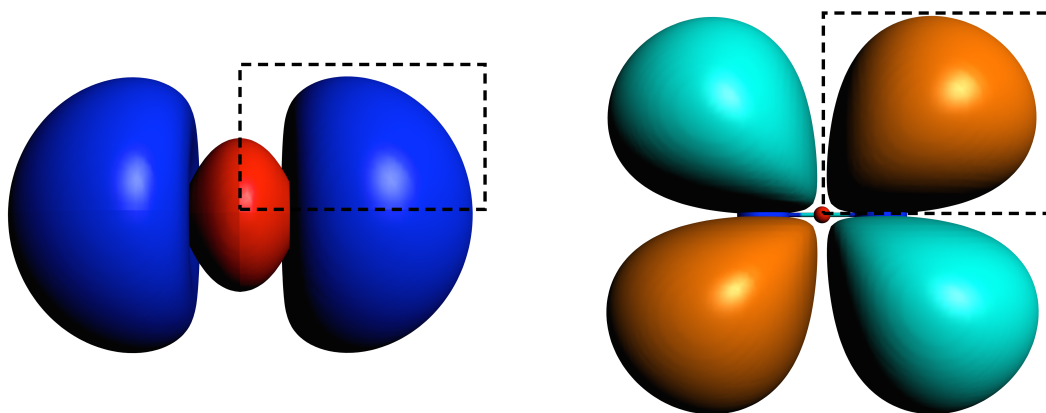


Figure 6.5: The HOMO (left) and LUMO (right) for a neutral nitrogen molecule. Black dashed boxes indicate the regions of the orbital plots that geometrically match the Fukui function plots in Figure 6.4.



Recently, it has not only been accepted that negative Fukui functions exist, but it is speculated that they may be useful in finding redox induced electron transfer (RIET) molecules [168–170]. A RIET molecule has the property that when the molecule as a whole is oxidized (removal of electrons), one of the atoms or functional groups in the molecule is reduced (gains electrons). Very few molecules have been shown to have the RIET property, since it is rare for the integration of  $\rho(\mathbf{r})$  over an entire atomic region to change in the direction opposite of the entire molecule. The same principle works for gradient bundles however, rather than just for atomic basins, and is much more common.

Regions of negative Fukui functions have been rationalized by orbital relaxation. When an electron is removed from a molecule, and the orbitals are allowed to relax, the orbitals that are still occupied contract toward the nucleus [174]. This causes electron density to actually increase in the nodal surfaces of the HOMO. The volume over which the Fukui function is negative varies from molecule to molecule. It has been postulated that molecules whose chemistry is not highly controlled by its frontier molecular orbitals will have more prominent negative Fukui function regions [170].

While the gradient bundle condensed Fukui functions calculated here match the frontier orbital picture, this may not necessarily always be true. The HOMO and LUMO neglect orbital relaxation and are only an approximation to where electron density will change most. A more accurate orbital picture of reactivity can be gained by calculating the frontier orbitals using the Fukui matrix. The Fukui matrix is defined as the derivative of the density matrix [106]. The density matrix can be expressed in terms of its basis functions,  $\varphi$ , as

$$P_{ij} = \langle \varphi_i(\mathbf{r}) | \rho(\mathbf{r}, \mathbf{r}') | \varphi_j(\mathbf{r}') \rangle \quad (6.5)$$

giving rise to the Fukui matrix

$$\mathbf{f} = \left[ f_{ij} = \left( \frac{\partial P_{ij}}{\partial N} \right)_{\nu(\mathbf{r})} \right] \quad (6.6)$$

where the Fukui function is given by the diagonal of  $\mathbf{f}$ . In general, the Fukui matrix is calculated using a finite difference just as the Fukui function is. The Fukui matrix can be diagonalized and decomposed as

$$\mathbf{f}^{\pm} = \sum_i \phi_i \mathbf{F}_{ii}^{\pm} \phi_i^T \quad (6.7)$$

where  $\mathbf{F}_{ii}$  are known as the Fukui eigenvalues and  $\phi_i$  are the Fukui orbitals.

For instances when the the frozen orbital approximation is completely accurate, there will be a single Fukui eigenvalue equal to one that completely represents the frontier molecular orbital. In general, one eigenvalue will be slightly less than one, and there will be other eigenfunctions that contribute to the actual regions where electron density is added or lost. Furthermore, Fukui eigenvalues can be negative, indicating orbitals that correspond to regions where negative Fukui functions are observed. We hypothesize that results from gradient bundle condensed Fukui functions will match the Fukui orbitals well, including negative regions. We plan on also extending this study to nonlinear molecules using the method advancements described in section 6.3.

## 6.2 Predicting Site Reactivity from the Ground State Charge Density

Calculating gradient bundle condensed Fukui functions has proven to be an indicator of chemical reactivity, but this method is not always computationally advantageous. When the electron count in a neutral system changes, electrons become unpaired and unrestricted calculations must be performed. These calculations are not always straightforward and often difficult to converge. Furthermore, from a theoretical standpoint they should not be necessary. The first theorem of DFT states that all properties are obtainable from the ground state charge density, without the need for a second calculation for the charged molecule. To this end, we are currently exploring ways of finding the reactive regions of a molecule from the ground state charge density alone.

Our method is to use differential geometry to observe the curvature and torsion of neighboring gradient bundles around each nuclei in a molecule. A related study was performed

calculating the curvature of atomic basin surfaces, in an attempt to associate the curvature of interatomic surfaces with properties such as electronegativity [175]. Any surface in 3D space can be characterized by its Gaussian curvature, which is the product of the two principle curvatures at a point on the surface. The total curvature of a surface is given by integration of the Gaussian curvature over the entire surface. Rather than directly calculating the curvature of ZFSs bounding gradient bundles (though this is also a path we plan to explore in future work), we are currently observing the effect of the differential geometry of the gradient vector field of  $\rho(\mathbf{r})$  that manifests in the size of gradient bundles. Specifically, we are calculating the cross-sectional area of gradient bundles at an isovalue of  $\rho(\mathbf{r})$ .

For linear molecules, this procedure is greatly simplified and can be approximated by simply measuring the width of rotational gradient bundles. Previous work plotting the Fukui function to predict reactivity has used the van der Waals radii of atoms as an isosurface cut-off with good results [176]. Figure 6.6 shows the width of gradient bundles around a nitrogen atom in  $\text{N}_2$  using the van der Waals radii defined at  $\rho(\mathbf{r}) = 0.002 \text{ e/bohr}^3$ . We find that electron density is most likely to be removed from the widest gradient bundles, i.e., the bundles with the largest surface area.

The rationale for this method is based on lower kinetic energy being an indicator of valence electrons as discussed in chapters 2 and 3. Electrons that are less confined (have lower density) will have lower kinetic energy. This idea can be seen in the simple Thomas-Fermi approximation of the kinetic energy, which is proportional to  $\rho(\mathbf{r})^{5/3}$  [177]. Gradient bundles have increased cross-sectional area when the charge density is decreasing rapidly from the center of the gradient bundle in a direction normal to the gradient paths in the bundle (i.e. have high Gaussian curvature). This is similar to calculating the sign of the Laplacian (which provides the direction the charge density is changing in the fastest), but goes beyond a local density approximation where only  $\rho(\mathbf{r})$  at a point is used. Instead, we are observing the more global property of the curvature of the charge density over surfaces which affects the volume of gradient bundles. Since electrons in larger gradient bundles

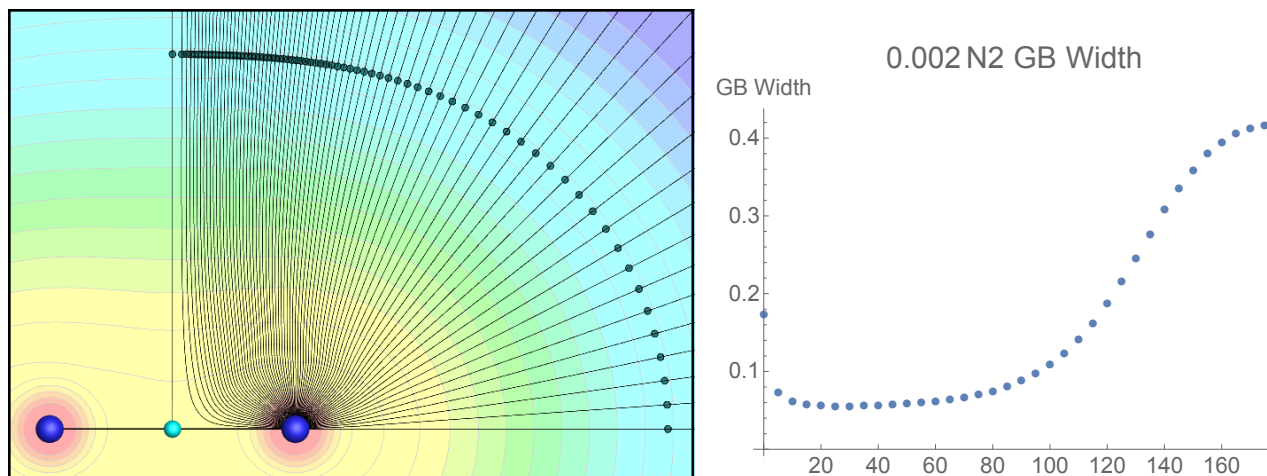


Figure 6.6: Left: A cutplane of rotational gradient bundles in N<sub>2</sub>. Black circles indicate points on the bounding zero-flux surfaces at an isovalue of  $\rho(\mathbf{r}) = 0.002$ . Right: The distance between GB points on the left, which we refer to as the width of the gradient bundles. The x-axis displays the gradient bundle angle, where 0° is the gradient bundle bounded by the interatomic surface and containing the bond path.

have more volume to occupy, they will move slower and have a lower average kinetic energy. Wider rotational gradient bundles indicate regions with low kinetic energy, and therefore can be used to locate valence electrons.

The widest gradient bundles in N<sub>2</sub> are pointing in the direction opposite the bond path, 180°, and there is a second maximum near the interatomic surface at 0°. This matches the two maxima in the gradient bundle condensed Fukui function and the HOMO for N<sub>2</sub>, as well as the location of valence electrons provided by analysis of the kinetic energy presented in chapter 2. Figure 6.7 shows the same analysis for F<sub>2</sub>. In this molecule there is a single maximum indicating the most energetically favorable region to remove electron density near 100°, exactly the same region that valence electrons were predicted to occupy using the kinetic energy density. Again, this matches the single lobe of the HOMO on each fluorine atom and the single maximum in  $f_{GB}^-$  in F<sub>2</sub>. The widest gradient bundles (at an isosurface of  $\rho(\mathbf{r})$  at the van der Waals radius) indicate the gradient bundles that will lose the most electron density when an electron is removed from the molecule as a whole.

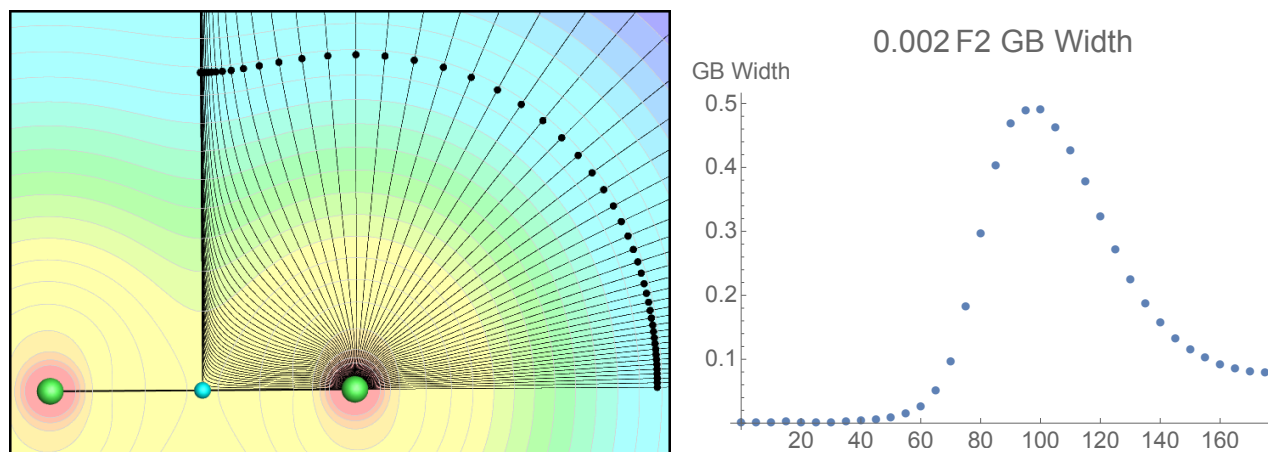


Figure 6.7: Left: A cutplane of rotational gradient bundles in  $F_2$ . Black circles indicate points on the bounding zero-flux surfaces at an isovalue of  $\rho(\mathbf{r}) = 0.002$ . Right: The width of the gradient bundles measured at the points marked in the left plot.

We are currently comparing the condensed Fukui function and gradient bundle sizes for 1-iodo-2-fluoro-ethyne, the molecule studied in chapter 4. ICCF has a more complex gradient bundle structure and bonding interactions than  $F_2$  or  $N_2$  and raises additional questions. Figure 6.8 shows the width of rotational gradient bundles at the van der Waals radii for each atom, as well as  $f_{GB,\%}^-$ . For the most part, maxima in the plots match up, showing that the size of gradient bundles is a good predictor of where electron density will be removed. However, the gradient bundle at  $0^\circ$  in  $C_2$  (along the  $C1-C2$  interatomic axis) does not fit this pattern. This gradient bundle is extremely large, yet its electron count does not decrease upon creating the cation of ICCF. In fact, this gradient bundle has a negative condensed Fukui function.

Further investigation of this gradient bundle in the  $C_2$  atomic basin is underway. The large difference in electronegativity between the fluorine and iodine atoms on the molecule may be what is causing the unusual gradient bundle to exist, since we do not observe the same feature in the gradient of the charge density for  $HCCH$ ,  $HCCI$  or  $HCCF$ .

We are also working on the more difficult task of predicting which gradient bundles will increase their electron counts when an electron is added to a molecule. How to find the

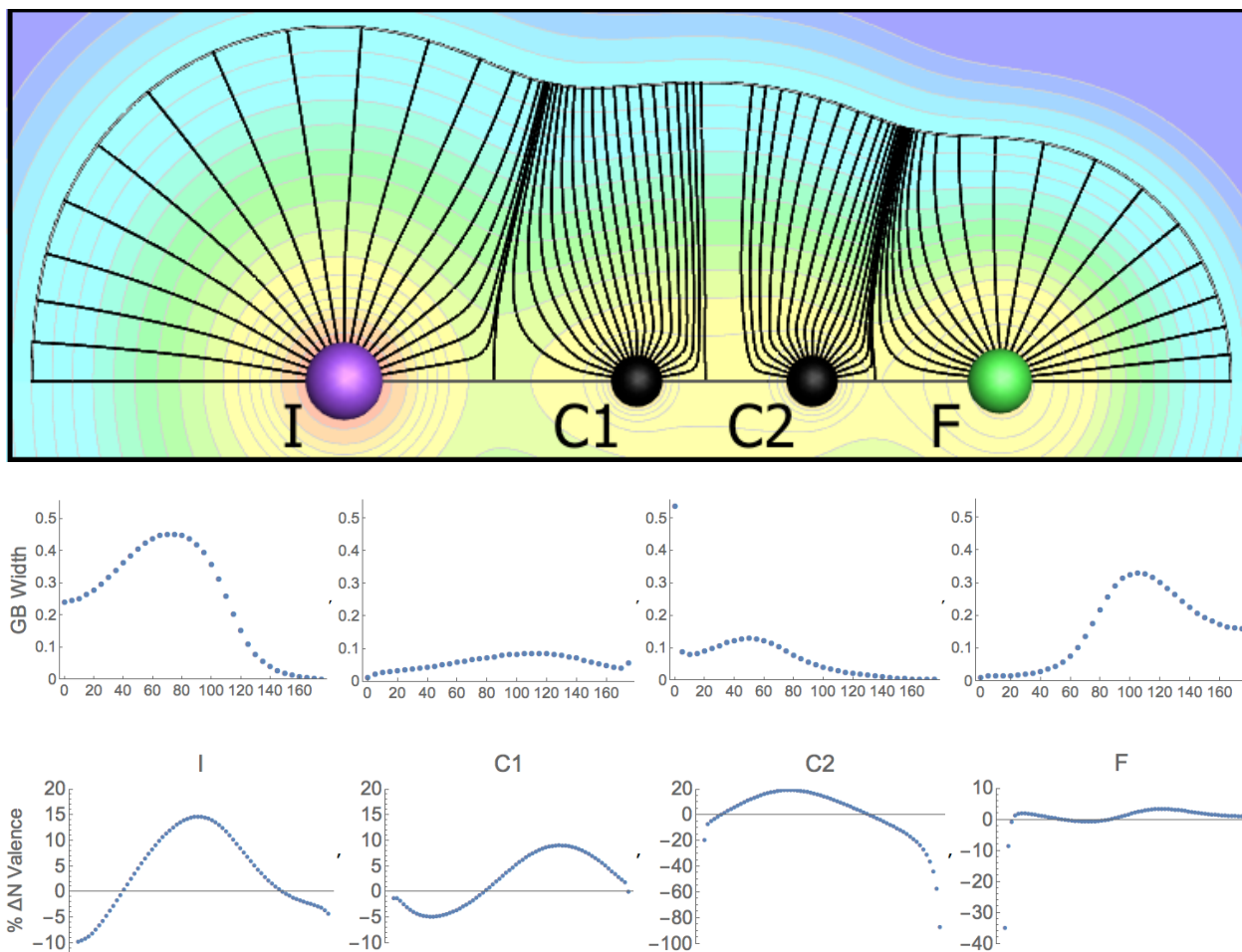


Figure 6.8: Top: A cut plane of ZFSs bounding rotational gradient bundles in ICCF. The isosurface of  $\rho(\mathbf{r}) = 0.002$  is traced with a black line to show where the width of gradient bundles is measured. Middle: The width of each GB at the van der Waals radius. Bottom:  $f_{GB,\%}^-$  for each atom in ICCF.

LUMO, so to say, in the ground state charge density is an open question in the field of CDFT. The ground state charge density is calculated as the sum of all occupied orbitals in the molecule, which includes the HOMO. The LUMO is unoccupied, making it difficult to understand where this orbital would be visible in the ground state charge density. However,  $\rho(\mathbf{r})$  contains all properties of a molecule, and therefore regions where it is most energetically favorable to add electrons must be identifiable using only  $\rho(\mathbf{r})$ . It is just a matter of discerning a method for capturing this property within the charge density structure.

We are currently investigating if a relationship exists between the LUMO and the rate at which gradient bundles are changing size as we move around each nuclei. Preliminary results show that electron density is added to one of two areas. One, the gradient bundles with the smallest cross-sectional areas match the LUMO. Or two, the regions with the steepest slope in the gradient bundle size plots, i.e. the gradient bundles that are changing size most quickly as we move radially around each nuclei. A full study of the differential geometry of the gradient bundles, including their integrated curvatures, may be necessary.

### 6.3 Gradient Bundle Analysis Method Improvements

It has been argued that studying the differential geometry of the full gradient vector field of  $\rho(\mathbf{r})$  would provide useful chemical information [32]. Bader stated that “further study of the gradient vector field of the electron density leads to a complete theory of structure...” [38]. While we have shown that studying the full gradient vector field provides additional information to standard QTAIM analysis, it is also apparent that this is a highly complex task from both a mathematical and computational standpoint. This is why our methods have only been applied to linear molecules. In order to apply GBA methods to more complex systems such as enzymes and metals, software and theoretical developments must be made.

The main obstacle in using gradient bundle analysis methods on non-linear systems is performing volume integrations. The program that we currently use to perform GBA, Tecplot [79], does not have a volume meshing algorithm. We are in the process of writing add-on packages for Tecplot that will enable volume integrations, rather than using our current

“trick” of multiplying properties in a molecular plane by  $2\pi r$ , where  $r$  is the distance from the internuclear axis (see section 2.4).

There are some recently developed programs that have algorithms for calculating atomic basin properties (see [178, 179] and references therein), but volume integration is still a major challenge in the entire field of QTAIM. Most of the available integration routines take advantage of the fact the the charge density and energy both decrease quickly as you move away from the nuclear CP, so most of the properties of interest are contained within the  $\beta$  sphere. The  $\beta$  sphere is defined as the largest sphere centered on a nuclear CP that does not extend past an interatomic surface in any direction. Therefore, it will be completely contained within the atomic basin but will generally not include parts of the basin near the bond critical points themselves. Since gradient bundles are not centered on nuclei,  $\beta$  sphere algorithms can not be utilized. Instead, a full volume meshing and quadrature integration method must be implemented.

Once volume integration of gradient bundles is possible, we will also be able to extend the work predicting site reactivity based on the size of gradient bundles. Rather than simply calculating the width of gradient bundles, we can compare the surface area of gradient bundles defined by the triangulation of a sphere around nuclear CPs as was shown in Figure 2.2. We plan to have one gradient bundle that contains the bond path of interest and then create a stereographic projection of gradient bundle properties (such as electron count and kinetic energy) centered around the bond path gradient bundle. This will enable us to see how gradient bundle properties vary in three-dimensions around a bonding interaction. Finally, we would like to test all of the gradient bundle methods described in this thesis using non-DFT methods. When these advancements to gradient bundle analysis are complete, we will be able to complete the projects discussed in the remaining sections of this chapter.

## 6.4 Applications to Aromaticity

The concept of aromaticity is often linked to electron delocalization. Electron delocalization is not an observable property however, meaning there is not a unique definition of



aromaticity [46]. Multiple theories try to describe and quantify delocalization of electrons based on the weighting of resonance structures [180], energy calculations from natural bonding orbitals [181], loge theory [182], and the Laplacian of the electron charge density [183], to name a few examples. There are a variety of aromaticity indicators based on these principles and others, such as the geometry-based harmonic oscillator model of aromaticity (HOMA) [184] and the magnetic-based nucleus-independent chemical shift (NICS) [185].

The use of aromaticity indicators based on QTAIM and ELF principles has been reviewed by Poater *et al.* [46]. The delocalization index (DI) (see section 1.1.4) between atoms in aromatic compounds [186], as well as the value of  $\rho(\mathbf{r})$  and  $\nabla\rho(\mathbf{r})$  at bond and ring CPs, has been found to often correlate to other aromaticity measures such as NICS and HOMA [187]. Values at  $(3, -1)$  CPs in the gradient field of ELF have similarly been used to quantify delocalization and hence aromaticity [188]. Santos *et al.* was able to better distinguish between many aromatic and anti-aromatic compounds using the  $\text{ELF}_\pi$  and  $\text{ELF}_\sigma$  values at CPs [47, 189]. However, when aromaticity values are compared using multiple indicators for the same set of molecules, it is not uncommon to have a different trend emerge depending on which indicators are used [190–192].

Poater *et al.* conclude their review on aromaticity by stating that, “the lack of universal nature of all indices available until now reinforces the stated belief of multidimensional character of aromaticity...” [46]. Studying the full gradient field of  $\rho(\mathbf{r})$ , as opposed to looking at DI between basins and values at critical points, may provide a method for capturing the multidimensional character ascribed to aromaticity. We are currently conducting a study comparing the gradient bundles of typical aromatic and anti-aromatic compounds including benzene, furan, cyclobuta-1,3-diene, and the anion and cation of cyclopenta-1,3-diene. We predict that the curvature of gradient paths terminating at ring CPs in cyclic aromatic compounds will be distinct between molecules that are aromatic and anti-aromatic in character.

## 6.5 Gradient Bundle Analysis in Closed Systems

The brittle behavior of face-centered cubic (FCC) iridium remains an unexplained property since it was first observed decades ago. Most FCC metals are highly ductile, whereas single crystal Ir is known to undergo brittle failure at room temperature and polycrystalline Ir undergoes intergranular fracture around 1000°C [193]. Many contradictory theories for the underlying cause of brittleness in Ir have been proposed including segregation of impurities at grain boundaries [194], bond directionality [195, 196], and ELF values [197, 198]. The search for an answer to this interesting behavior remains an area of active investigation [199].

The metallurgical community believes that the brittle behavior of iridium can be explained using the electron charge density. We have studied the values of  $\rho(\mathbf{r})$ ,  $\nabla\rho(\mathbf{r})$ , and directionality at bond and ring CPs in FCC Ir and Pt. As in studies of ELF values at these points, the charge density values between the two metals are similar, yet their observed properties (including Cauchy pressure) are distinct. Additionally, the full topology of the charge density, including all the edges of irreducible bundles, are identical between Ir and Pt. However, we believe there may be some distinguishing features in the full gradient field of  $\rho(\mathbf{r})$ .

To test this hypothesis, we are performing full gradient bundle analysis on Pt and Ir, using both cluster calculations performed within ADF and full periodic structures. We are currently qualitatively comparing the gradient fields in planes that should be important to determination of the shear moduli and hence Cauchy pressure in these materials. Once the advancements to GBA proposed in section 6.3 are complete, we will also quantitatively compare 3D gradient paths and volumes of gradient bundles in Ir and Pt, with the goal of discerning the unique properties of these materials despite their apparent physical similarities in  $\rho(\mathbf{r})$ .

## 6.6 Electrostatic Preorganization and Active Site Model Size

The origin of the catalytic power of enzymes is a matter of debate. A popular theory is that the active site of the enzyme binds strongly to the transition state of the substrate molecule, which lowers the activation energy of the reaction. It is also often stated that the catalytic groups in the active site are perfectly oriented to facilitate the reaction. Some chemists attribute the catalytic efficiency of enzymes to transition state stabilization (TSS) [200] while others attempt to explain the catalytic rates based on reactant state destabilization (RSD) [201]. All of these proposals are contingent on the method used to define the reference state used for comparison.

Warshel has argued that electrostatic preorganization is the main contributor to the high catalytic ability of enzymes [127, 202]. When a reaction is catalyzed in a polar solution such as water, the solvent molecules around the solute are randomly oriented with dipole vectors pointing in various directions. As the reaction proceeds through the transition state, the water dipoles must orient themselves in such a way as to facilitate the lowering of the activation barrier. This rearrangement of the electron density costs energy (due to water–water interactions being broken in order for hydrogen bonds to the substrate to be formed that stabilize the TS) and is contained in the activation energy of the transition state. In an enzyme, the active site is preorganized, meaning that the electrostatic environment is already oriented to stabilize the transition state. Dipole vectors are properly arranged through the positioning of charged and polar residues as well as the orientation of any water molecules in the active site. There is little or zero energy cost for organizing the electron density in the active site, since it is already in the most favorable position.

Capturing the electrostatic preorganization in a computer simulation requires taking into account the entire protein and solvent environment. However, this does not necessarily mean one must calculate the charge density of the entire protein to a high degree of accuracy. The electron density has been described as near-sighted, meaning that the density in one region is not highly dependent on the charge density further away [203]. The electron

density based on a calculation including only the atoms in the active site should closely match the density calculated using the full enzyme, as long as the nuclear positions are provided from a full enzyme calculation. Therefore, it should be possible to capture the electrostatic preorganization in the active site of an enzyme just by performing a single point QM calculation on the active site, keeping in mind that the nuclear coordinates must first be calculated using full protein simulations.

We are currently designing a computational experiment to see if this hypothesis is valid. There are two goals of this study. The first is to see if we do indeed capture the electrostatic preorganization of the active site just by running QM calculations on the nuclear coordinates from a full protein simulation (without any electrostatic embedding). The second task is to see how large of an active site model needs to be modeled in order to capture the correct electron charge density distribution. De Proft recently performed a study [204] showing that the Hirshfeld charges on a neutral residue in a protein converge very quickly as the active site model size changes (from a 3 to 11 Å radii around a central atom). However, the atomic charges did not converge as well when the residue was charged. This study only looked at one system (human 2-cysteine peroxiredoxin thioredoxin peroxidase B) and reported the convergence of atomic charges using the Hirshfeld, Hirshfeld-I, and natural population analysis (NPA) charges.

The nuclear coordinates from QM/DMD calculations for both the carboxypeptidase A (CPA) and histone deacetylase 8 (HDAC8) enzymes presented in chapter 5 will be used in this study. Following the methods from [204], active site models for QM single point calculations are being created using a 3, 5, 7, 9, and 11 Å radius around the metal cations. Any amino acid residue that has at least one atom within this cut-off will be completely retained, with dangling residues being capped with hydrogen atoms. Additionally, each of the five size models for each enzyme will be run as a gas-phase calculation, using COSMO, and with point charges to simulate the electrostatic environment from the full protein, for a total of thirty calculations. The charge densities will be analyzed using QTAIM and

gradient bundle analysis. The bader charges of the metal ions, substrate carbonyls, and water molecule (which is a key component of the reaction mechanism in both of the enzymes tested) will be compared for each calculation. The full gradient field of the charge density will be compared by overlaying gradient paths and comparing the stereographic projections of the amount of charge density in gradient bundles around key bond paths as described in section 6.3.

The results of this study may help to determine the level of computational cost required in order to properly describe the electrostatic environment in the active site of a metalloenzyme. The size of the active site proved to be an important factor in correctly calculating the energetics of reaction in HDAC8, so it will be interesting to see if this is also important for determining the electron density distribution, or if the theory of the near-sightedness of  $\rho(\mathbf{r})$  holds. Initial testing of electrostatic embedding (using point charges) did not qualitatively change the reaction profile in HDAC8, but it did change the exact energies at stationary points. Here we will see if electrostatic embedding or implicit solvation affects the charge density calculated in the active site of enzymes compared to purely gas phase calculations. Following this study, we will also focus on further quantifying the way changes in  $\rho(\mathbf{r})$  equate to changes in activation energies and reaction rates.

## 6.7 Conclusions

In this thesis I have demonstrated that gradient bundle analysis is a useful and meaningful method for understanding and predicting chemical properties. Partitioning the electron charge density into gradient bundles allows for distributions of well-defined properties to be calculated within atomic basins and bond bundles. The structure of the valence electrons in molecules—bonding and lone pair regions—can be determined using gradient bundles, highlighting regions that are most likely to undergo reaction. This structure can be used to define kinetic energy bonding regions, whose valence electron counts correlate to experimentally determined bond dissociation energies in diatomic molecules. The motion of zero-flux surfaces bounding gradient bundles can be used to explain chemical reactivity, reproducing

results obtained using standard frontier orbital theory, but without the need to decompose the charge density into orbital contributions. We are currently working on enhancing the predictive capabilities of GBA, allowing for the most reactive sites on a molecule to be located using only the ground state charge density.

Advancements to gradient bundle analysis highlighted in this chapter will enable quantitative chemical bonding analysis of non-symmetric molecules. The motivation for this work is to improve our ability to rationally design materials and enzymes with novel properties. Current bonding models can not always fully explain the relationships between the structure of molecules and solids with their observed properties, such as our lack of understanding the brittle behavior of iridium described in section 6.5. Rational enzyme design methods could also benefit from the use of gradient bundle analysis, as we showed in section 5.3 that more traditional methods for analyzing the charge density, such as QTAIM, can not always connect the topology of  $\rho(\mathbf{r})$  in active sites to predicted and experimental activation energies. Using gradient bundles to analyze the full topology and geometry of the electron charge density will enable a higher resolution model of chemical bonding to be created, allowing for improved prediction of the reactivity in designed enzymes, materials, and molecules.

## REFERENCES CITED

- [1] R. F. W. Bader. *Atoms in Molecules: A quantum theory*. Clarendon Press, Oxford, UK, 1990.
- [2] P. Geerlings, F. De Proft, and W. Langenaeker. Conceptual density functional theory. *Chemical Reviews*, 103(5):1793–1874, 2003.
- [3] Heidi K. Privett, Gert Kiss, Toni M. Lee, Rebecca Blomberg, Roberto A. Chica, Leonard M. Thomas, Donald Hilvert, Kendall N. Houk, and Stephen L. Mayo. Iterative approach to computational enzyme design. *Proceedings of the National Academy of Sciences*, 109(10):3790–3795, 2012.
- [4] Gert Kiss, Nihan Çelebi-Ölçüm, Rocco Moretti, David Baker, and K. N. Houk. Computational enzyme design. *Angewandte Chemie International Edition*, 52(22):5700–5725, 2013.
- [5] Maria P. Frushicheva, Matthew J. L. Mills, Patrick Schopf, Manoj K. Singh, Ram B. Prasad, and Arieh Warshel. Computer aided enzyme design and catalytic concepts. *Current Opinion in Chemical Biology*, 21:56–62, 2014.
- [6] Alexandre Zanghellini. de novo computational enzyme design. *Current Opinion in Biotechnology*, 29:132–138, 2014.
- [7] Kenneth B. Wiberg, Richard F. W. Bader, and Clement D. H. Lau. Theoretical analysis of hydrocarbon properties. 2. additivity of group properties and the origin of strain energy. *Journal of the American Chemical Society*, 109(4):1001–1012, 1987.
- [8] P. L. A. Popelier. Quantum molecular similarity. 1. bcp space. *The Journal of Physical Chemistry A*, 103(15):2883–2890, 1999.
- [9] Chérif F. Matta, Norberto Castillo, and Russell J. Boyd. Extended weak bonding interactions in dna:  $\pi$ -stacking (base-base), base-backbone, and backbone-backbone interactions. *The Journal of Physical Chemistry B*, 110(1):563–578, 2006.
- [10] C. F. Matta and R. F. W. Bader. Atoms-in-molecules study of the genetically encoded amino acids. III. Bond and atomic properties and their correlations with experiment including mutation-induced changes in protein stability and genetic coding. *Proteins*, 52(3):360–399, 2003.

- [11] U. A. Chaudry, , and P. L. A. Popelier. Estimation of pka using quantum topological molecular similarity descriptors: Application to carboxylic acids, anilines and phenols. *The Journal of Organic Chemistry*, 69(2):233–241, 2004.
- [12] P. F. Zou and R. F. W. Bader. A topological definition of a wigner–seitz cell and the atomic scattering factor. *Acta Crystallographica Section A*, 50(6):714–725, Nov 1994.
- [13] A. Martín Pendás, Aurora Costales, and Víctor Luaña. Ions in crystals: The topology of the electron density in ionic materials. i. Fundamentals. *Physical Review B*, 55: 4275–4284, Feb 1997.
- [14] Mark Eberhart. Charge-density-shear-moduli relationships in aluminum-lithium alloys. *Physical Review Letters*, 87:205503, Oct 2001.
- [15] Richard F. W. Bader. Private communication, April 1987.
- [16] Shant Shahbazian and Mansour Zahedi. Letter to the editor: The concept of chemical bond - some like it fuzzy but others concrete. *Foundations of Chemistry*, 9(1):85–95, 2006.
- [17] Jordi Poater, Miquel Solà, and F. Matthias Bickelhaupt. A model of the chemical bond must be rooted in quantum mechanics, provide insight, and possess predictive power. *Chemistry – A European Journal*, 12(10):2902–2905, 2006.
- [18] Kenneth G. Dyall and Knut Fægri Jr. Finite nucleus effects on relativistic energy corrections. *Chemical Physics Letters*, 201(1–4):27–32, 1993.
- [19] Dirk Andrae. Finite nuclear charge density distributions in electronic structure calculations for atoms and molecules. *Physics Reports*, 336(6):413–525, 2000.
- [20] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler. Chemistry with adf. *Journal of Computational Chemistry*, 22(9):931–967, 2001.
- [21] P. L. A. Popelier. Characterization of a dihydrogen bond on the basis of the electron density. *The Journal of Physical Chemistry A*, 102(10):1873–1878, 1998.
- [22] C. F. Matta and R. J. Boyd. *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*. Wiley-VCH, Verlag GmbH and Co. KGaA, Weinheim, 2007.
- [23] Sławomir Janusz Grabowski. Ab initio calculations on conventional and unconventional hydrogen bonds—study of the hydrogen bond strength. *The Journal of Physical Chemistry A*, 105(47):10739–10746, 2001.



- [24] Thomas A. Ford. An ab initio study of the properties of some lithium-bonded complexes - comparison with their hydrogen-bonded analogues. 2. Natural bond orbital and quantum theory of atoms in molecules analysis. *Computational and Theoretical Chemistry*, 1042:63–68, 2014.
- [25] M.E. Eberhart. The metallic bond: Elastic properties. *Acta Materialia*, 44(6):2495–2504, 1996.
- [26] Mark E. Eberhart and Travis E. Jones. Cauchy pressure and the generalized bonding model for nonmagnetic bcc transition metals. *Physical Review B*, 86:134106, Oct 2012.
- [27] M.E. Eberhart and A.F. Giamei. The visualization and use of electronic structure for metallurgical applications. *Materials Science and Engineering: A*, 248(1–2):287–295, 1998.
- [28] X. F. Wang, T. E. Jones, Y. Wu, Z. P. Lu, S. Halas, T. Durakiewicz, and M. E. Eberhart. An electronic criterion for assessing intrinsic brittleness of metallic glasses. *The Journal of Chemical Physics*, 141(2):024503, 2014.
- [29] Alexander Popkov and Martin Breza. Why is monoalkylation versus bis-alkylation of the Ni(II) complex of the schiff base of (S)-N-(2-benzoylphenyl)-1-benzylpyrrolidine-2-carboxamide and glycine so selective? MP2 modelling and topological QTAIM analysis of chiral metallocomplex synthons of  $\alpha$ -amino acids used for the preparation of radiopharmaceuticals for positron emission tomography. *Journal of Radioanalytical and Nuclear Chemistry*, 286(3):829–833, 2010.
- [30] Samantha Jenkins, Albeiro Restrepo, Jorge David, Dulin Yin, and Steven R. Kirk. Spanning QTAIM topology phase diagrams of water isomers  $w_4$ ,  $w_5$ , and  $w_6$ . *Physical Chemistry Chemical Physics*, 13:11644–11656, 2011.
- [31] G. R. Runtz, R. F. W. Bader, and R. R. Messer. Definition of bond paths and bond directions in terms of the molecular charge distribution. *Canadian Journal of Chemistry*, 55(16):3040–3045, 1977.
- [32] Paul L. A. Popelier. *Atoms in Molecules: An introduction*. Pearson Education Limited, Essex, England, 2000.
- [33] R. F. W. Bader. Atoms in molecules. *Accounts of Chemical Research*, 18(1):9–15, 1985.
- [34] Jonathan Miorelli, Tim Wilson, Amanda Morgenstern, Travis Jones, and Mark E. Eberhart. A full topological analysis of unstable and metastable bond critical points. *ChemPhysChem*, 16(1):152–159, 2015.

- [35] Arne Haaland, Dimitry J. Shorokhov, and Natalya V. Tverdova. Topological analysis of electron densities: Is the presence of an atomic interaction line in an equilibrium geometry a sufficient condition for the existence of a chemical bond? *Chemistry – A European Journal*, 10(18):4416–4421, 2004.
- [36] Arne Haaland, Dimitry J. Shorokhov, and Natalya V. Tverdova. Topological analysis of electron densities: Is the presence of an atomic interaction line in an equilibrium geometry a sufficient condition for the existence of a chemical bond? *Chemistry – A European Journal*, 10(24):6210–6210, 2004.
- [37] Cina Foroutan-Nejad, Shant Shahbazian, and Radek Marek. Toward a consistent interpretation of the QTAIM: Tortuous link between chemical bonds, interactions, and bond/line paths. *Chemistry – A European Journal*, 20(32):10140–10152, 2014.
- [38] Richard Frederick William Bader, Paul Lode Albert Popelier, and Todd Alan Keith. Theoretical definition of a functional group and the molecular orbital paradigm. *Angewandte Chemie International Edition in English*, 33(6):620–631, 1994.
- [39] James S. M. Anderson, Paul W. Ayers, and Juan I. Rodriguez Hernandez. How ambiguous is the local kinetic energy? *The Journal of Physical Chemistry A*, 114(33):8884–8895, 2010.
- [40] Leon Cohen. Local kinetic energy in quantum mechanics. *The Journal of Chemical Physics*, 70(2):788–789, 1979.
- [41] Swapan K. Ghosh and Robert G. Parr. Density-determined orthonormal orbital approach to atomic energy functionals. *The Journal of Chemical Physics*, 82(7):3307–3315, 1985.
- [42] Chérif F. Matta and Jesús Hernández-Trujillos. Bonding in polycyclic aromatic hydrocarbons in terms of the electron density and of electron delocalization. *The Journal of Physical Chemistry A*, 107(38):7496–7504, 2003.
- [43] A. D. Becke and K. E. Edgecombe. A simple measure of electron localization in atomic and molecular systems. *The Journal of Chemical Physics*, 92(9):5397–5403, 1990.
- [44] Andreas Savin, Reinhard Nesper, Steffen Wengert, and Thomas F. Fässler. ELF: The electron localization function. *Angewandte Chemie International Edition in English*, 36(17):1808–1832, 1997.
- [45] Andreas Savin, Ove Jepsen, Jürgen Flad, Ole Krogh Andersen, Heinzwerner Preuss, and Hans Georg von Schnering. Electron localization in solid-state structures of the elements: the diamond structure. *Angewandte Chemie International Edition in English*, 31(2):187–188, 1992.

- [46] Jordi Poater, Miquel Duran, Miquel Solà, and Bernard Silvi. Theoretical evaluation of electron delocalization in aromatic molecules by means of atoms in molecules (AIM) and electron localization function (ELF) topological approaches. *Chemical Reviews*, 105(10):3911–3947, 2005.
- [47] J. C. Santos, W. Tiznado, R. Contreras, and P. Fuentealba. Sigma–pi separation of the electron localization function and aromaticity. *The Journal of Chemical Physics*, 120(4):1670–1673, 2004.
- [48] M. Kohout. A measure of electron localizability. *International Journal of Quantum Chemistry*, 97(1):651–658, 2004.
- [49] Stephan N. Steinmann, Yirong Mo, and Clemence Corminboeuf. How do electron localization functions describe  $\pi$ -electron delocalization. *Physical Chemistry Chemical Physics*, 13:20584–20592, 2011.
- [50] Frank R. Wagner, Viktor Bezugly, Miroslav Kohout, and Yuri Grin. Charge decomposition analysis of the electron localizability indicator: A bridge between the orbital and direct space representation of the chemical bond. *Chemistry – A European Journal*, 13(20):5724–5741, 2007.
- [51] H. L. Schmider and A. D. Becke. Chemical content of the kinetic energy density. *Journal of Molecular Structure: THEOCHEM*, 527:51–61, 2000.
- [52] Heiko Jacobsen. Topology maps of bond descriptors based on the kinetic energy density and the essence of chemical bonding. *Physical Chemistry Chemical Physics*, 15:5057–5066, 2013.
- [53] Paul W. Ayers, Robert G. Parr, and Agnes Nagy. Local kinetic energy and local temperature in the density-functional theory of electronic structure. *International Journal of Quantum Chemistry*, 90:309–326, 2002.
- [54] T. Bitter, K. Ruedenberg, and W. H. E. Schwarz. Toward a physical understanding of electron-sharing two-center bonds. I. General aspects. *Journal of Computational Chemistry*, 28(1):411–422, 2007.
- [55] Yang Yang. Hexacoordinate bonding and aromaticity in silicon phthalocyanide. *Journal of Physical Chemistry A*, 114:13257–13267, 2010.
- [56] Mark Eberhart. A quantum description of the chemical bond. *Philosophical Magazine Part B*, 81(8):721–729, 2001.
- [57] Jason Miller. *Relative Critical Sets in  $\mathbb{R}^n$  and Applications to Image Analysis*. PhD thesis, Chapel Hill, 1998.

- [58] James Damon. Generic structure of two-dimensional images under gaussian blurring. *SIAM Journal on Applied Mathematics*, 59(1):97–138, 1998.
- [59] D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach. Ridges for image analysis. *Journal of Mathematical Imaging and Vision*, 4:353–373, 1994.
- [60] Travis E. Jones and Mark E. Eberhart. The irreducible bundle: Further structure in the kinetic energy distribution. *Journal of Chemical Physics*, 130:204108, 2009.
- [61] Travis E. Jones, Mark E. Eberhart, Scott Imlay, and Craig Mackey. Bond bundles and the origins of functionality. *Journal of Physical Chemistry A*, 115:12582–12585, 2011.
- [62] Travis E. Jones. Nucleophilic substitution: A charge density perspective. *Journal of Physical Chemistry A*, 116:4233–4237, 2012.
- [63] Joel L. Wolk, Tova Hoz, Harold Basch, and Shmaryahu Hoz. Quantification of the various contributors to rate enhancement in nucleophilic strain releasing reactions. *Journal of Organic Chemistry*, 66:915–918, 2001.
- [64] Travis E. Jones, Mark E. Eberhart, Scott Imlay, Craig Mackey, and Greg B. Olson. Better alloys with quantum design. *Physical Review Letters*, 109:125506, 2012.
- [65] Felix Latourte, Xiaoding Wei, Zechariah D. Feinberg, Alban de Vaucorbeil, Phuong Tran, Gregory B. Olson, and Horacio D. Espinosa. Design and identification of high performance steel alloys for structures subjected to underwater impulsive loading. *International Journal of Solids and Structure*, 49(13):1573–1587, 2012.
- [66] T.E. Jones and M.E. Eberhart. The bond bundle in open systems. *International Journal of Quantum Chemistry*, 110:1500–1505, 2010.
- [67] Mark E. Eberhart and Travis E. Jones. The two faces of chemistry: can they be reconciled? *Foundations of Chemistry*, 15(3):277–285, 2013.
- [68] C. Fonseca Guerra, J. G. Snijders, G. te Velde, and E. J. Baerends. Towards an order-N DFT method. *Theoretical Chemistry Accounts*, 99(6):391–403, 1998.
- [69] John P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33:8822–8824, 1986.
- [70] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, 1996.

- [71] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11):6671–6687, 1992.
- [72] Juan I. Rodríguez, Paul W. Ayers, Andreas W. Götz, and F. L. Castillo-Alvarado. Virial theorem in the kohn-sham density-functional theory formalism: Accurate calculation of the atomic quantum theory of atoms in molecules energies. *Journal of Chemical Physics*, 131:21101, 2009.
- [73] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin-density calculations- a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.
- [74] P. J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry.*, 98(45):11623–11627, 1994.
- [75] Y. Zhao and D.G. Truhlar. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and non covalent interactions. *Journal of Chemical Physics*, 125:194101, 2006.
- [76] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37:785–789, Jan 1988.
- [77] Travis E. Jones, Mark E. Eberhart, and Dennis P. Clougherty. Topology of the spin-polarized charge density in bcc and fcc iron. *Physical Review Letters*, 100(1):017208, Jan 2008.
- [78] Christian Carbogno, Axel Gross, Jörg Meyer, and Karsten Reuter. O<sub>2</sub> adsorption dynamics at metal surfaces: Non-adiabatic effects, dissociation and dissipation. In Ricardo Díez Muiño and Heriberto Fabio Busnengo, editors, *Dynamics of Gas-Surface Interactions*, volume 50 of *Springer Series in Surface Sciences*, pages 389–419. Springer Berlin Heidelberg, 2013.
- [79] Tecplot. <http://www.tecplot.com/>, 2014.
- [80] M. H. Stans and Lewis M. Branscomb. *Bond Dissociation Energies in Simple Molecules*. NSRDS, <http://www.nist.gov/data/nsrds/NSRDS-NBS31.pdf>, 1970.
- [81] C. Schweitzer and R. Schmidt. Physical mechanisms of generation and deactivation of singlet oxygen. *Chemical Reviews*, 103:1685–1757, 2003.

- [82] Leon Cohen. Representable local kinetic energy. *The Journal of Chemical Physics*, 80(9):4277–4279, 1984.
- [83] A.Martín Pendás, Evelio Francisco, MiguelA. Blanco, and Carlo Gatti. Bond paths as privileged exchange channels. *Chemistry – A European Journal*, 13(33):9362–9371, 2007.
- [84] J. C. Slater. The virial and molecular structure. *Journal of Chemical Physics*, 1: 687–691, 1933.
- [85] Klaus Ruedenberg. The physical nature of the chemical bond. *Reviews of Moderns Physics*, 34(2):326–376, 1962.
- [86] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Clarendon Press, Oxford, 1989.
- [87] Amanda Morgenstern, Tim Wilson, Jonathan Miorelli, Travis Jones, and M. E. Eberhart. In search of an intrinsic chemical bond. *Computational and Theoretical Chemistry*, 1053:31–37, 2015.
- [88] Kenichi Fukui, Teijiro Yonezawa, and Haruo Shingu. A molecular orbital theory of reactivity in aromatic hydrocarbons. *The Journal of Chemical Physics*, 20(4):722–725, 1952.
- [89] William Ogilvy Kermack and Robert Robinson. An explanation of the property of induced polarity of atoms and an interpretation of the theory of partial valencies on an electronic basis. *Journal of Chemical Society, Transactions*, 121:427, 1922.
- [90] Richard F. W. Bader, T. Tung NguyenDang, and Yoram Tal. Quantum topology of molecular charge distributions. ii. molecular structure and its change. *The Journal of Chemical Physics*, 70(9):4316–4329, 1979.
- [91] Alfredo Guevara-García, Eleonora Echegaray, Alejandro Toro-Labbe, Samantha Jenkins, Steven R. Kirk, and Paul W. Ayers. Pointing the way to the products? Comparison of the stress tensor and the second-derivative tensor of the electron density. *Journal of Chemical Physics*, 134:234106, 2011.
- [92] Paul W. Ayers and Samantha Jenkins. An electron-preceding perspective on the deformation of materials. *Journal of Chemical Physics*, 130(15):154104, 2009.
- [93] A. Guevara-García, P. W. Ayers, S. Jenkins, S. R. Kirk, E. Echegaray, and A. Toro-Labbe. Electronic stress as a guiding force for chemical bonding. *Topics in Current Chemistry*, 351:103–124, 2014.

- [94] Robert G. Parr, Robert A. Donnelly, Mel Levy, and William E. Palke. Electronegativity: The density functional viewpoint. *Journal of Chemical Physics*, 68:3801–3807, 1978.
- [95] Robert G. Parr and Ralph G. Pearson. Absolute hardness: Companion parameter to absolute electronegativity. *Journal of American Chemical Society*, 105:7512–7516, 1983.
- [96] Naba K. Ray, Leonard Samuels, and Robert G. Parr. Studies of electronegativity equalization. *Journal of Chemical Physics*, 70:3680–3684, 1979.
- [97] Peter Politzer and Harel Weinstein. Some relations between electronic distribution and electronegativity. *Journal of Chemical Physics*, 71:4218–4220, 1979.
- [98] Vincent Tognetti, Christophe Morell, and Laurent Joubert. Atomic electronegativities in molecules. *Chemical Physics Letters*, 635:111–115, 2015.
- [99] Paul W. Ayers and Robert G. Parr. Local hardness equalization: Exploiting the ambiguity. *Journal of Chemical Physics*, 128:184108, 2008.
- [100] P. W. Ayers and R. G. Parr. Variational principles for describing chemical reactions: The Fukui function and chemical hardness revisited. *Journal of American Chemical Society*, 122:2010, 2000.
- [101] S. K. Ghosh. Energy derivatives in density-functional theory. *Chemical Physics Letters*, 172:77, 1990.
- [102] S. K. Ghosh and M. Berkowitz. A classical fluid-like approach to density-functional formalism of many-electron systems. *Journal of Chemical Physics*, 83:2976, 1985.
- [103] R. T. Sanderson. Partial charges on atoms in organic compounds. *Science*, 121:207–208, 1955.
- [104] H. Chermette. Chemical reactivity indexes in density functional theory. *Journal of Computational Chemistry*, 20:129–154, 1999.
- [105] Robert G. Parr and Weitao Yang. Density functional approach to the frontier-electron theory of chemical reactivity. *Journal of American Chemical Society*, 106:4049–4050, 1984.
- [106] Patrick Bultinck, Dorien Clarisse, Paul W. Ayers, and Ramon Carbó-Dorca. The Fukui matrix: a simple approach to the analysis of the Fukui function and its positive character. *Physical Chemistry Chemical Physics*, 13:6110–6115, 2011.

- [107] Patrick Bultinck, Dimitri Van Neck, Guillaume Acke, and Paul W. Ayers. Influence of electron correlation and degeneracy on the Fukui matrix and extension of frontier molecular orbital theory to correlated quantum chemical methods. *Physical Chemistry Chemical Physics*, 14:2408–2416, 2012.
- [108] Patrick Bultinck, Carlos Cardenas, Patricio Fuentealba, Paul A. Johnson, and Paul W. Ayers. How to compute the Fukui matrix and function for systems with (quasi-) degenerate states. *Journal of Chemical Theory and Computation*, 10:202–210, 2014.
- [109] Paul W. Ayers, Robert C. Morrison, and Ram K. Roy. Variational principles for describing chemical reactions: Condensed reactivity indices. *Journal of Chemical Physics*, 116:8731–8744, 2002.
- [110] Patrick Bultinck, Stijn Fias, Christian Van Alsenoy, Paul W. Ayers, and Ramon Carbó-Dorca. Critical thoughts on computing atom condensed fukui functions. *Journal of Chemical Physics*, 127:034102, 2007.
- [111] Jerzy Cioslowski, Martin Martinov, and Stacey T. Mixon. Atomic fukui indices from the topological theory of atoms in molecules applied to Hartree-Fock and correlated electron densities. *Journal of Physical Chemistry*, 97:10948–10951, 1993.
- [112] Renato R. Contreras, Patricio Fuentealba, Marcelo Galván, and Patricia Pérez. A direct evaluation of regional Fukui functions in molecules. *Chemical Physics Letters*, 304:405–413, 1999.
- [113] S. Arulmozhiraja and P. Kolandaivel. Condensed Fukui function: dependency on atomic charges. *Molecular Physics*, 90(1):55–62, 1997.
- [114] F. De Proft, C. Van Alsenoy, A. Peeters, W. Langenaeker, and P. Geerlings. Atomic charges, dipole moments, and Fukui functions using the hirshfeld partitioning of the electron density. *Journal of Computational Chemistry*, 23:1198–2002, 2002.
- [115] ISO/TC 12. Iso 80000-s:2009(en) quantities and unites – part 2: Mathematical signs and symbols to be used in the natural sciences and technology, 2009.
- [116] E. van Lenthe and E.J. Baerends. Optimized slater-type basis sets for the elements 1-118. *Journal of Computational Chemistry*, 24:1142, 2003.
- [117] E. van Lenthe, E. J. Baerends, and J. G. Snijders. Relativistic regular two-component hamiltonians. *Journal of Chemical Physics*, 99:4597–4610, 1993.
- [118] E. van Lenthe, E. J. Baerends, and J. G. Snijders. Relativistic total energy using regular approximations. *Journal of Chemical Physics*, 101:9783–9792, 1994.



- [119] E. van Lenthe, A. E. Ehlers, and E. J. Baerends. Geometry optimization in the zero order regular approximation for relativistic effects. *Journal of Chemical Physics*, 110: 8943–8953, 1999.
- [120] J.I. Rodriguez, R.F.W. Bader, P.W. Ayers, C. Michel, A.W. Götz, and C. Bo. A high performance grid-based algorithm for computing QTAIM properties. *Chemical Physics Letters*, 472:149, 2009.
- [121] J.I. Rodriguez. An efficient method for computing the QTAIM topology of a scalar field: The electron density case. *Journal of Computational Chemistry*, 34:681, 2013.
- [122] Bernard Meunier, Samuël P. de Visser, and Sason Shaik. Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. *Chemical Reviews*, 104(9):3947–3980, 2004.
- [123] F. Albert Cotton, Edward E. Hazen, and Margaret J. Legg. Staphylococcal nuclease: Proposed mechanism of action based on structure of enzyme—thymidine 3,5-bisphosphate—calcium ion complex at 1.5-Å resolution. *Proceedings of the National Academy of Sciences*, 76(6):2551–2555, 1979.
- [124] Stephen W. Ragsdale. *The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment*, chapter Biochemistry of Methyl-Coenzyme M Reductase: The Nickel Metalloenzyme that Catalyzes the Final Step in Synthesis and the First Step in Anaerobic Oxidation of the Greenhouse Gas Methane, pages 125–145. Springer Netherlands, Dordrecht, 2014.
- [125] Meng Zhao, Hai-Bo Wang, Liang-Nian Ji, and Zong-Wan Mao. Insights into metalloenzyme microenvironments: biomimetic metal complexes with a functional second coordination sphere. *Chemical Society Reviews*, 42:8360–8375, 2013.
- [126] Robert R. Knowles and Eric N. Jacobsen. Attractive noncovalent interactions in asymmetric catalysis: Links between enzymes and small molecule catalysts. *Proceedings of the National Academy of Sciences*, 107(48):20678–20685, 2010.
- [127] Arie Warshel. Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *Journal of Biological Chemistry*, 273(42):27035–27038, 1998.
- [128] Arie Warshel. Computer simulations of enzyme catalysis: Methods, progress, and insights. *Annual Review of Biophysics and Biomolecular Structure*, 32(1):425–443, 2003.

- [129] Esteban Gabriel Vega-Hissi, Rodrigo Tosso, Ricardo Daniel Enriz, and Lucas Joel Gutierrez. Molecular insight into the interaction mechanisms of amino-2h-imidazole derivatives with bace1 protease: A qm/mm and qtain study. *International Journal of Quantum Chemistry*, 115(6):389–397, 2015.
- [130] Rodrigo D. Tosso, Sebastian A. Andujar, Lucas Gutierrez, Emilio Angelina, Ricaurte Rodríguez, Manuel Nogueras, Héctor Baldoni, Fernando D. Suvire, Justo Cobo, and Ricardo D. Enriz. Molecular modeling study of dihydrofolate reductase inhibitors. molecular dynamics simulations, quantum mechanical calculations, and experimental corroboration. *Journal of Chemical Information and Modeling*, 53(8):2018–2032, 2013.
- [131] Emilio L. Angelina, Sebastián A. Andujar, Rodrigo D. Tosso, Ricardo D. Enriz, and Nélida M. Peruchena. Non-covalent interactions in receptor–ligand complexes. a study based on the electron charge density. *Journal of Physical Organic Chemistry*, 27(2): 128–134, 2014.
- [132] Zhong Ni, Rongzhong Jin, Huayou Chen, and Xianfu Lin. Just an additional hydrogen bond can dramatically reduce the catalytic activity of bacillus subtilis lipase a I12T mutant: An integration of computational modeling and experimental analysis. *Computers in Biology and Medicine*, 43(11):1882–1888, 2013.
- [133] Manuel Sparta, David Shirvanyants, Feng Ding, Nikolay V. Dokholyan, and Anastassia N. Alexandrova. Hybrid dynamics simulation engine for metalloproteins. *Biophysical Journal*, 103(4):767–776, 2012.
- [134] Nikolay V. Dokholyan, Sergey V. Buldyrev, H Eugene Stanley, and Eugene I. Shakhnovich. Discrete molecular dynamics studies of the folding of a protein-like model. *Folding and Design*, 3(6):577–587, 11 1998.
- [135] Nikolay V Dokholyan. Studies of folding and misfolding using simplified models. *Current Opinion in Structural Biology*, 16(1):79–85, 2 2006.
- [136] Feng Ding, Weihua Guo, Nikolay V. Dokholyan, Eugene I. Shakhnovich, and Joan-Emma Shea. Reconstruction of the src-sh3 protein domain transition state ensemble using multiscale molecular dynamics simulations. *Journal of Molecular Biology*, 350 (5):1035–1050, 7 2005.
- [137] Feng Ding, Douglas Tsao, Huifen Nie, and Nikolay V. Dokholyan. Ab initio folding of proteins with all-atom discrete molecular dynamics. *Structure*, 16(7):1010–1018, 7 2008.
- [138] Manuel Sparta and Anastassia N Alexandrova. How metal substitution affects the enzymatic activity of catechol-o-methyltransferase. *PLoS ONE*, 7(10):e47172, 2012.

- [139] Crystal E. Valdez and Anastassia N. Alexandrova. Why urease is a di-nickel enzyme whereas the ccra -lactamase is a di-zinc enzyme. *The Journal of Physical Chemistry B*, 116(35):10649–10656, 09 2012.
- [140] Manuel Sparta, Crystal E. Valdez, and Anastassia N. Alexandrova. Metal-dependent activity of fe and ni acireductone dioxygenases: How two electrons reroute the catalytic pathway. *Journal of Molecular Biology*, 425(16):3007–3018, 8 2013.
- [141] Silvia Álvarez Santos, Àngels González-Lafont, Jose’ M. Lluch, Baldomero Oliva, and Francesc X. Avile’s. Theoretical study of the role of arginine 127 in the water-promoted mechanism of peptide cleavage by carboxypeptidase a. *New Journal of Chemistry*, 22: 319–326, 1998.
- [142] Alexandra Vardi-Kilshtain, Gil Shoham, and Amiram Goldblum. Anhydride formation is not a valid mechanism for peptide cleavage by carboxypeptidase-a: a semiempirical reaction pathway study. *Molecular Physics*, 101(17):2715–2724, 2003.
- [143] Michelle W.Y. Szeto, Jon I. Mujika, Jolanta Zurek, Adrian J. Mulholland, and Jeremy N. Harvey. QM/MM study on the mechanism of peptide hydrolysis by carboxypeptidase A. *Journal of Molecular Structure:THEOCHEM*, 898(1–3):106–114, 2009.
- [144] Hidong Kim and William N. Lipscomb. Crystal structure of the complex of carboxypeptidase a with a strongly bound phosphonate in a new crystalline form: comparison with structures of other complexes. *Biochemistry*, 29(23):5546–5555, 1990.
- [145] Reinhart Ahlrichs, Michael Bär, Marco Häser, Hans Horn, and Christoph Kölmel. Electronic structure calculations on workstation computers: The program system turbomole. *Chemical Physics Letters*, 162(3):165–169, 1989.
- [146] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45:13244–13249, Jun 1992.
- [147] Jianmin Tao, John P. Perdew, Viktor N. Staroverov, and Gustavo E. Scuseria. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, 91:146401, Sep 2003.
- [148] Ansgar Schäfer, Hans Horn, and Reinhart Ahlrichs. Fully optimized contracted gaussian basis sets for atoms Li to Kr. *The Journal of Chemical Physics*, 97(4):2571–2577, 1992.
- [149] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7:3297–3305, 2005.

- [150] A. Klamt and G. Schuurmann. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of Chemical Society, Perkin Transactions 2*, pages 799–805, 1993.
- [151] *ADF2014*. SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2014.
- [152] Michael Seth and Tom Ziegler. Range-separated exchange functionals with slater-type functions. *Journal of Chemical Theory and Computation*, 8(3):901–907, 2012.
- [153] C. Cory Pye and Tom Ziegler. An implementation of the conductor-like screening model of solvation within the amsterdam density functional package. *Theoretical Chemistry Accounts*, 101(6):396–408, 1999.
- [154] Michael Haberland, Rusty L. Montgomery, and Eric N. Olson. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature Reviews Genetics*, 10(1):32–42, 01 2009.
- [155] Jessica E. Bolden, Melissa J. Peart, and Ricky W. Johnstone. Anticancer activities of histone deacetylase inhibitors. *Nature Reviews Drug Discovery*, 5(9):769–784, 09 2006.
- [156] Paul A Marks and Ronald Breslow. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nature Biotechnology*, 25(1):84–90, 01 2007.
- [157] James E Bradner, Nathan West, Melissa L Grachan, Edward F Greenberg, Stephen J Haggarty, Tandy Warnow, and Ralph Mazitschek. Chemical phylogenetics of histone deacetylases. *Nature Chemical Biology*, 6(3):238–243, 03 2010.
- [158] Marielle Paris, Marina Porcelloni, Monica Binaschi, and Daniela Fattori. Histone deacetylase inhibitors: From bench to clinic. *Journal of Medicinal Chemistry*, 51(6):1505–1529, 2008.
- [159] Nobuaki Suzuki, Takayoshi Suzuki, Yosuke Ota, Tatsuya Nakano, Masaaki Kurihara, Haruhiro Okuda, Takao Yamori, Hiroki Tsumoto, Hidehiko Nakagawa, and Naoki Miyata. Design, synthesis, and biological activity of boronic acid-based histone deacetylase inhibitors. *Journal of Medicinal Chemistry*, 52(9):2909–2922, 2009.
- [160] Stephanie L. Gantt, Samuel G. Gattis, , and Carol A. Fierke. Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion. *Biochemistry*, 45(19):6170–6178, 2006.

- [161] Ruibo Wu, Shenglong Wang, Nengjie Zhou, Zexing Cao, and Yingkai Zhang. A proton-shuttle reaction mechanism for histone deacetylase 8 and the catalytic role of metal ions. *Journal of the American Chemical Society*, 132(27):9471–9479, 2010.
- [162] Michael R. Nechay, Nathan M. Gallup, Amanda Morgenstern, Quentin A. Smith, Mark E. Eberhart, and Anastassia N. Alexandrova. Histone deacetylase 8: Characterization of physiological divalent metal catalysis. *Journal of Physical Chemistry B*, 2016. DOI: 10.1021/acs.jpcb.6b00997.
- [163] Alessandro Vannini, Cinzia Volpari, Paola Gallinari, Philip Jones, Marco Mattu, Andrea Carfi, Raffaele De Francesco, Christian Steinkühler, and Stefania Di Marco. Substrate binding to histone deacetylases as shown by the crystal structure of the hdac8–substrate complex. *EMBO reports*, 8(9):879–884, 2007.
- [164] Stefan Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25(12):1463–1473, 2004.
- [165] Analytical and Biological Products. *Dojindo.com Metal Chelates*, pp 252-253.
- [166] Viktor N. Staroverov, Gustavo E. Scuseria, Jianmin Tao, and John P. Perdew. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *The Journal of Chemical Physics*, 119(23):12129–12137, 2003.
- [167] Clémence Corminboeuf, Po Hu, Mark E. Tuckerman, and Yingkai Zhang. Unexpected deacetylation mechanism suggested by a density functional theory qm/mm study of histone-deacetylase-like protein. *Journal of the American Chemical Society*, 128(14):4530–4531, 2006.
- [168] Paul W. Ayers. Can one oxidize an atom by reducing the molecule that contains it? *Physical Chemistry Chemical Physics*, 8:3387–3390, 2006.
- [169] Junia Melin, Paul W. Ayers, and Joseph Vincent Ortiz. Removing electrons can increase the electron density: a computational study of negative fukui functions. *The Journal of Physical Chemistry A*, 111(40):10017–10019, 2007.
- [170] Eleonora Echegaray, Carlos Cárdenas, Sandra Rabi, Nataly Rabi, Sungmin Lee, Farnaz Heidar Zadeh, Alejandro Toro-Labbe, James S. M. Anderson, and Paul W. Ayers. In pursuit of negative fukui functions: examples where the highest occupied molecular orbital fails to dominate the chemical reactivity. *Journal of Molecular Modeling*, 19(7):2779–2783, 2012.

- [171] Felipe A. Bulat, Eduardo Chamorro, Patricio Fuentealba, and Alejandro Toro-Labbe. Condensation of frontier molecular orbital fukui functions. *The Journal of Physical Chemistry A*, 108(2):342–349, 2004.
- [172] P. Fuentealba, P. Pérez, and R. Contreras. On the condensed fukui function. *The Journal of Chemical Physics*, 113(7):2544–2551, 2000.
- [173] Ram Kinkar Roy, Sourav Pal, and Kimihiko Hirao. On non-negativity of fukui function indices. *The Journal of Chemical Physics*, 110(17):8236–8245, 1999.
- [174] Paul W. Ayers, Robert C. Morrison, and Ram K. Roy. Variational principles for describing chemical reactions: Condensed reactivity indices. *The Journal of Chemical Physics*, 116(20):8731–8744, 2002.
- [175] Paul L.A. Popelier. On the differential geometry of interatomic surfaces. *Canadian Journal of Chemistry*, 74(6):829–838, 1996.
- [176] Libero J. Bartolotti and Paul W. Ayers. An example where orbital relaxation is an important contribution to the fukui function. *The Journal of Physical Chemistry A*, 109(6):1146–1151, 2005.
- [177] N.H. March. *Theory of the Inhomogeneous Electron Gas*. Plenum Press, 1983.
- [178] Paul L. A. Popelier. Fully analytical integration over the 3d volume bounded by the sphere in topological atoms. *The Journal of Physical Chemistry A*, 115(45):13169–13179, 2011.
- [179] M. Rafat and P. L. A. Popelier. Visualization and integration of quantum topological atoms by spatial discretization into finite elements. *Journal of Computational Chemistry*, 28(16):2602–2617, 2007.
- [180] Ramon M. Parrondo, Padeleimon Karafiloglou, Rafael R. Pappalardo, and Enrique Sanchez Marcos. Calculation of the weights of resonance structures of molecules in solution. *The Journal of Physical Chemistry*, 99(17):6461–6467, 1995.
- [181] Maria Grazia Giuffreda, Maurizio Bruschi, and Hans Peter Lüthi. Electron delocalization in linearly  $\pi$ -conjugated systems: A concept for quantitative analysis. *Chemistry – A European Journal*, 10(22):5671–5680, 2004.
- [182] R. Daudel, R. F. W. Bader, M. E. Stephens, and D. S. Borrett. The electron pair in chemistry. *Canadian Journal of Chemistry*, 52(8):1310–1320, 1974.

- [183] Cina Foroutan-Nejad, Zahra Badri, Shant Shahbazian, and Parviz Rashidi-Ranjbar. The laplacian of electron density versus  $\text{nics}_{zz}$  scan: Measuring magnetic aromaticity among molecules with different atom types. *The Journal of Physical Chemistry A*, 115(45):12708–12714, 2011.
- [184] Tadeusz Marek Krygowski. Crystallographic studies of inter- and intramolecular interactions reflected in aromatic character of  $\pi$ -electron systems. *Journal of Chemical Information and Computer Sciences*, 33(1):70–78, 1993.
- [185] Paul von Ragué Schleyer, Christoph Maerker, Alk Dransfeld, Haijun Jiao, and Nicolaas J. R. van Eikema Hommes. Nucleus-independent chemical shifts: a simple and efficient aromaticity probe. *Journal of the American Chemical Society*, 118(26):6317–6318, 1996.
- [186] Richard F. W. Bader, Andrew Streitwieser, Arnst Neuhaus, Keith E. Laidig, , and Peter Speers. Electron delocalization and the fermi hole. *Journal of the American Chemical Society*, 118(21):4959–4965, 1996.
- [187] S. T. Howard and T. M. Krygowski. Benzenoid hydrocarbon aromaticity in terms of charge density descriptors. *Canadian Journal of Chemistry*, 75(9):1174–1181, 1997.
- [188] A. Savin, B. Silvi, and F. Coionna. Topological analysis of the electron localization function applied to delocalized bonds. *Canadian Journal of Chemistry*, 74(6):1088–1096, 1996.
- [189] Juan C. Santos, Juan Andres, Arie Aizman, and Patricio Fuentealba. An aromaticity scale based on the topological analysis of the electron localization function including and contributions. *Journal of Chemical Theory and Computation*, 1(1):83–86, 2005.
- [190] Tadeusz M. Krygowski, Krzysztof Ejsmont, Beata T. Stepień, Michał K. Cyrański, Jordi Poater, and Miquel Solà. Relation between the substituent effect and aromaticity. *The Journal of Organic Chemistry*, 69(20):6634–6640, 2004.
- [191] Jordi Poater, Isidoro Garcia-Cruz, Francesc Illas, and Miquel Sola. Discrepancy between common local aromaticity measures in a series of carbazole derivatives. *Physical Chemistry Chemical Physics*, 6:314–318, 2004.
- [192] Michał K. Cyrański, Tadeusz M. Krygowski, Alan R. Katritzky, and Paul von Ragué Schleyer. To what extent can aromaticity be defined uniquely?. *The Journal of Organic Chemistry*, 67(4):1333–1338, 2002.
- [193] S. S. Hecker, D. L. Rohr, and D. F. Stein. Brittle fracture in iridium. *Metallurgical Transactions A*, 9(4):481–488, 1978.

- [194] M.A Fortes and B Ralph. A field-ion microscope study of segregation to grain boundaries in iridium. *Acta Metallurgica*, 15(5):707–720, 1967.
- [195] Yu. N. Gornostyrev, M. I. Katsnelson, N. I. Medvedeva, O. N. Mryasov, A. J. Freeman, and A. V. Trefilov. Peculiarities of defect structure and mechanical properties of iridium: Results of *ab initio* electronic structure calculations. *Physical Review B*, 62: 7802–7808, Sep 2000.
- [196] H.R. Gong. Ideal mechanical strength and interface cohesion property of ir-base superalloys from first principles calculation. *Materials Chemistry and Physics*, 126(1–2): 284–288, 2011.
- [197] Sami Kamran, Kuiying Chen, Liang Chen, and Linruo Zhao. Electronic origin of anomalously high shear modulus and intrinsic brittleness of fcc ir. *Journal of Physics: Condensed Matter*, 20(8):085221, 2008.
- [198] M. Černý and J. Pokluda. The theoretical shear strength of fcc crystals under superimposed triaxial stress. *Acta Materialia*, 58(8):3117–3123, 5 2010.
- [199] C.P. Liang, G.H. Li, and H.R. Gong. Concerning the brittleness of iridium: An elastic and electronic view. *Materials Chemistry and Physics*, 133(1):140–143, 2012.
- [200] L. Pauling. Molecular architecture and biological reactions. *Chemical and Engineering News*, 24:1375–1377, 1946.
- [201] W. P. Jencks. *Catalysis in Chemistry and Enzymology*. Dover Publication, New York, 1986.
- [202] Arie Warshel, Pankaz K. Sharma, Mitsunori Kato, Yun Xiang, Hanbin Liu, and Mats H. M. Olsson. Electrostatic basis for enzyme catalysis. *Chemical Reviews*, 106 (8):3210–3235, 2006.
- [203] E. Prodan and W. Kohn. Nearsightedness of electronic matter. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33):11635–11638, 2005.
- [204] Danny E. P. Vanpoucke, Julianna Oláh, Frank De Proft, Veronique Van Speybroeck, and Goedeke Roos. Convergence of atomic charges with the size of the enzymatic environment. *Journal of Chemical Information and Modeling*, 55(3):564–571, 2015.



## APPENDIX A

### FUNCTIONAL DEPENDENCE

Plots of the average kinetic energy per electron in gradient bundles for  $N_2$  using various DFT methods are presented in Figure A.1. We have tested LDA, BLYP, B3LYP, and M06L functionals, in addition to the PBE data presented in chapter 2. All plots are geometrically very similar. Local maxima and minima occur in the same gradient bundles (maxima at  $80^\circ$ , minima at  $0^\circ$  and  $180^\circ$ ), and current investigation into the curvatures of the plots at local maxima and minima suggests these plots not to be significantly different.

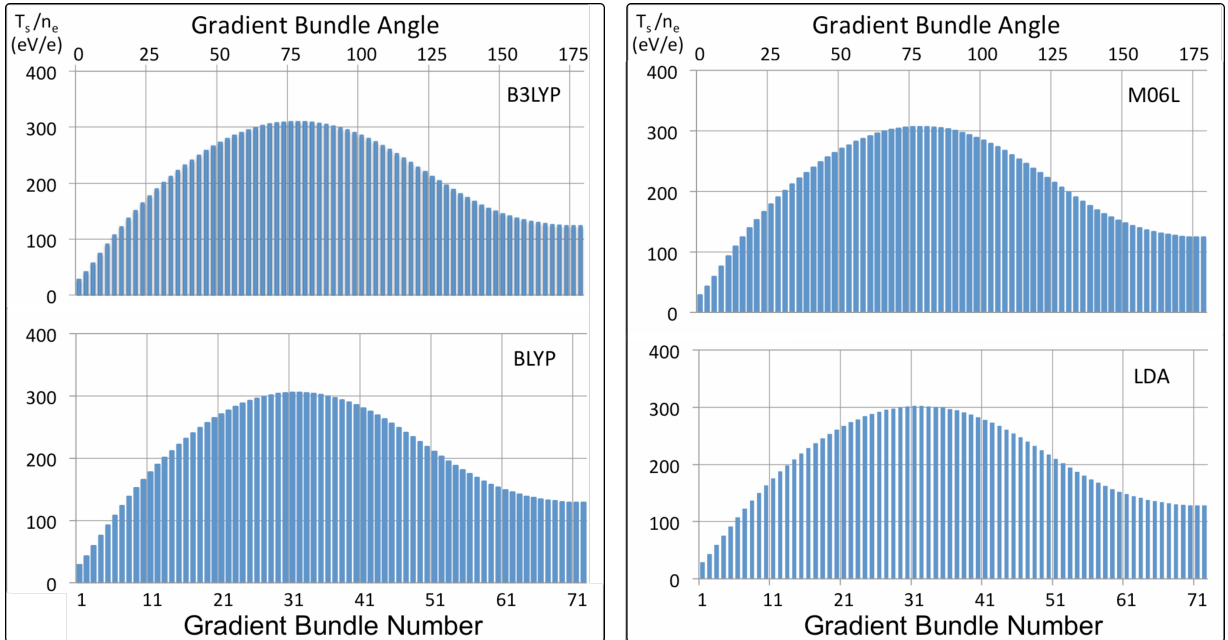


Figure A.1: Average KE/e in gradient bundles seeded every  $2.5^\circ$  around a nucleus in  $N_2$  using LDA, BLYP, B3LYP, and M06L functionals within the ADF software package. Gradient bundle angle is measured with respect to the bond path. Single point calculations were run using a TZP basis set using the geometry obtained from initial PBE calculations.

## APPENDIX B

### HDAC8 METAL SWAPPING DATA

Table B.1 provides experimental binding energies from [165] for various metal ions in DTPA and GEDTA used for metal swapping calculations for HDAC8. Table B.2 lists calculated  $\Delta G$  values for metal swapping between DTPA and GEDTA, using  $\text{Co}^{2+}$  as a reference point. Table B.3 contains the experimental  $\Delta G$  values for the same metal swapping reactions. Table B.4 shows the difference between experimental and calculated values of  $\Delta G$  for metal swapping.

Table B.1: List of literature binding energies with regards to a series of chelating agents for all relevant metal ions studied in HDAC8.

Metal	DTPA		GEDTA	
	Stability constant	$\Delta G$ binding ( <i>kcal/mol</i> )	Stability constant	$\Delta G$ binding ( <i>kcal/mol</i> )
$\text{Co}^{2+}$	18.4	-25.	12.5	-17.
$\text{Fe}^{2+}$	16.55	-23.	11.92	-16.
$\text{Mg}^{2+}$	9.3	-13.	5.21	-7.1
$\text{Mn}^{2+}$	15.6	-21.	12.3	-17.
$\text{Ni}^{2+}$	20.32	-28.	13.6	-19.
$\text{Zn}^{2+}$	18.75	-26.	14.5	-20.

Table B.2: Calculated  $\Delta G$  values for DTPA metal swapped with GEDTA (*kcal/mol*).

	Co <sup>2+</sup>	Fe <sup>2+</sup>	Mg <sup>2+</sup>	Mn <sup>2+</sup>	Ni <sup>2+</sup>	Zn <sup>2+</sup>
Co <sup>2+</sup>	0	-2.25	-3.41	-1.96	1.31	-2.37
Fe <sup>2+</sup>		0	-1.16	0.29	3.56	-0.12
Mg <sup>2+</sup>			0	1.45	4.72	1.03
Mn <sup>2+</sup>				0	3.27	-0.41
Ni <sup>2+</sup>					0	-3.68
Zn <sup>2+</sup>						0

Table B.3: Experimental  $\Delta G$  values for DTPA metal swapped with GEDTA (*kcal/mol*).

	Co <sup>2+</sup>	Fe <sup>2+</sup>	Mg <sup>2+</sup>	Mn <sup>2+</sup>	Ni <sup>2+</sup>	Zn <sup>2+</sup>
Co <sup>2+</sup>	0	-1.73	-2.47	-3.55	1.12	-2.25
Fe <sup>2+</sup>		0	-0.74	-1.81	2.85	-0.52
Mg <sup>2+</sup>			0	-1.08	3.59	0.22
Mn <sup>2+</sup>				0	4.66	1.30
Ni <sup>2+</sup>					0	-3.37
Zn <sup>2+</sup>						0

Table B.4: Calculated  $\Delta G$  – experimental  $\Delta G$  values for DTPA metal swapped with GEDTA (*kcal/mol*).

	Co <sup>2+</sup>	Fe <sup>2+</sup>	Mg <sup>2+</sup>	Mn <sup>2+</sup>	Ni <sup>2+</sup>	Zn <sup>2+</sup>
Co <sup>2+</sup>	0	-0.52	-0.94	1.59	0.19	-0.12
Fe <sup>2+</sup>		0	-0.42	2.11	0.71	0.40
Mg <sup>2+</sup>			0	2.53	1.13	0.82
Mn <sup>2+</sup>				0	-1.40	-1.71
Ni <sup>2+</sup>					0	-0.31
Zn <sup>2+</sup>						0

## APPENDIX C

### PERMISSIONS

This appendix provides permissions from journals and co-authors of published material presented in this dissertation.

#### C.1 Journal Permissions

Excerpted from the licenses, copyright, and permissions of Royal Society of Chemistry journals, including *Physical Chemistry Chemical Physics* and *Physica Scripta*: “You do not need to request permission to reuse your own figures, diagrams, etc, that were originally published in a Royal Society of Chemistry publication. However, permission should be requested for use of the whole article or chapter except if reusing it in a thesis. If you are including an article or book chapter published by us in your thesis please ensure that your co-authors are aware of this.”

Excerpted from the permissions of Elsevier journals, including *Computational and Theoretical Chemistry*: “Authors can include their articles in full or in part in a thesis or dissertation for non-commercial purposes.”

Excerpted from the permissions of American Chemical Society journals, including *Journal of Physical Chemistry B*: “ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).”

#### C.2 Author Permissions

Tim Wilson (Chapters 2 and 4), twilson@mines.edu

I’m sending this email indicating my authorization for our co-authored papers, “In Search of an Intrinsic Chemical Bond” and “The Influence of Zero-Flux Surface Motion on Chemical

Reactivity” to appear in your dissertation.

Sincerely,

Tim Wilson

Jonathan Miorelli (Chapters 2 and 4), jmiorell@mines.edu

You have my permission to use both, “In Search of an Intrinsic Chemical Bond” and “The Influence of Zero-Flux Surface Motion on Chemical Reactivity” in your dissertation.

Jonathan Miorelli

Travis Jones (Chapter 2), trjones@mines.edu

You have my permission to use our paper “In Search of Intrinsic Chemical Bond” in your dissertation.

And congratulations!

Travis

Charles Morgenstern (Chapter 4), cmorgens@mines.edu

You have my permission to include “The Influence of Zero-Flux Surface Motion on Chemical Reactivity” in your thesis.

Charles Morgenstern

Crystal Valdez (Chapter 5), cediev87@gmail.com

I give you permission to use parts of our co-authored submitted paper, “Predictive methods for computational metalloenzyme redesign - a test case with carboxypeptidase A,” in any parts of your dissertation.

Best,

Crystal

Quentin Smith (Chapter 5), schrodinator@gmail.com

No problem.

Nathan Gallup (Chapter 5), gallup@chem.ucla.edu

Yes, you may use it.

Michael Nechay (Chapter 5), michaelnechay@gmail.com

I grant you my permission to use parts of our co-authored published paper “Histone Deacetylase 8: Characterization of Physiological Divalent Metal Catalysis” in your dissertation.

-Michael Nechay